

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 10, 2005, 21:23:50 ; Search time 329.5 Seconds  
(without alignments)  
5748.378 Million cell updates/sec

Title: US-10-054-683-18

Perfect score: 4796

Sequence: 1 catctgcacttccaactgc.....actgttacatttcaaaaaa 2640

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 4004546

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2 1/USP0 spool p/US10054683/runat 10012005 174509 174509/app query.fasta\_1.2823  
-DB-A Geneseq 23Sep04 -OPMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USBR=US10054683 @CGN 1 1 398 @runat 10012005 174509 15490 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 23Sep04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3984	83.1	734	2	AAR87037
2	3984	83.1	734	6	ABJ19246 Human PH3
3	3984	83.1	734	7	ADB75186 Human can
4	3969.5	82.8	735	7	ADB75186 Prostate
5	3900	81.3	753	7	ADB75182 Prostate
6	3551	74.0	651	2	AAR87034 Human PH3
7	2440	50.9	735	2	AAR87036 Mouse PH3
8	1530	31.9	820	6	ABR39696 Human met
9	1524.5	31.8	787	6	AAU83636 Human PRO
10	1524.5	31.8	787	6	ABU80783 Human PRO

11	1524.5	31.8	787	6	ABO33749	Novel hum
12	1524.5	31.8	787	6	ABU82092	Novel hum
13	1524.5	31.8	787	6	ABJ72272	Human PRO
14	1524.5	31.8	787	6	ABJ72400	Human PRO
15	1524.5	31.8	787	6	ABO34295	Human sec
16	1524.5	31.8	787	7	ABJ72102	Human mem
17	1524.5	31.8	787	7	ADB83580	Novel hum
18	1524.5	31.8	787	7	ADB80686	Novel hum
19	1524.5	31.8	787	7	ADB73227	Novel hum
20	1524.5	31.8	787	7	ADB78309	Novel hum
21	1524.5	31.8	787	7	ADB84957	Human PRO
22	1524.5	31.8	787	7	ADB78063	Novel hum
23	1524.5	31.8	787	7	ADB87129	Human PRO
24	1524.5	31.8	787	7	ADB84711	Human PRO
25	1524.5	31.8	787	7	ADB83826	Novel hum
26	1524.5	31.8	787	7	ADB72981	Novel hum
27	1524.5	31.8	787	7	ADC36819	Human PRO
28	1524.5	31.8	787	7	ADC21809	Human PRO
29	1524.5	31.8	787	7	ADC49840	Novel hum
30	1524.5	31.8	787	7	ADC49039	Novel hum
31	1524.5	31.8	787	7	ADC49556	Novel hum
32	1524.5	31.8	787	7	ADC47417	Novel hum
33	1524.5	31.8	787	7	ADC47162	Novel hum
34	1524.5	31.8	787	7	ADC78037	Novel hum
35	1524.5	31.8	787	7	ADD06272	Novel hum
36	1524.5	31.8	787	7	ADC77791	Novel hum
37	1524.5	31.8	787	7	ADD50754	Novel hum
38	1524.5	31.8	787	7	ADD51000	Novel hum
39	1524.5	31.8	787	7	ADD50481	Human PRO
40	1524.5	31.8	787	7	ADD50235	Human PRO
41	1524.5	31.8	787	7	ADD51246	Novel hum
42	1524.5	31.8	787	8	ADC48793	Novel hum
43	1524.5	31.8	787	8	ADE20964	Novel hum
44	1524.5	31.8	787	8	ADE05808	Human PRO
45	1524.5	31.8	787	8	ADD75037	Human PRO

ALIGNMENTS

RESULT 1

AAR87037

ID AAR87037 standard; protein; 734 AA.

XX AAR87037;

AC AAR87037;

XX 30-SEP-1996 (first entry)

DT Human PH30 beta chain sperm protein.

DE Human PH30 beta chain sperm protein.

XX Human PH30 beta chain sperm protein; contraceptive;

KW FEE integrin binding domain.

XX Homo sapiens.

OS Homo sapiens.

XX Key

FT Key

FT Region

XX WO9535118-A1.

XX 28-DEC-1995.

XX 06-JUN-1995; 95WO-US007295.

XX 20-JUN-1994; 94US-00264101.

XX (MERI ) MERCK & CO INC.

XX PI Alves K, Gupta SK, Hollis GF;

XX WPI; 1996-058212/06.

XX N-FSDB; AAT07328.

XX

PT Human and mouse sperm protein PH30 beta chain and related DNA - useful in  
 XX contraceptive vaccines.

XX Example 2; Page 45-48; 85pp; English.

XX Human PH30 beta chain sperm protein, having a FEE integrin binding  
 CC domain, is 58.9% identical to mouse and 56.5% identical to guinea pig  
 CC PH30 beta. The protein may be produced recombinantly and used in a  
 CC contraceptive composition containing an effective adjuvant and an amount  
 CC of sperm protein which is effective for the stimulation of antibodies  
 CC which bind to sperm protein in vivo, thereby preventing or substantially  
 CC reducing the rate of sperm-egg fusion

XX Sequence 734 AA;

#### Alignment Scores:

Pred. No.:	0	Length:	734
Score:	3984.00	Matches:	734
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	83.07%	Indels:	0
DB:	2	Gaps:	0

US-10-054-683-18 (1-2640) x AAR87037 (1-734)

Qy	72	ATGTGGGTCTTGTCTCTCAGCGGCTCGCGGGCTCGGATGGACAGTAATTTTGTAT	131
Db	1	MetTrpValLeuPheLeuLeuSerGlyLeuGlyLeuArgMetAspSerAsnPheAsp	20
Qy	132	AGTTTACCTGTGCAAAATTTACAGTTCCGGAGAAATACGGTCAATTAATAAGGAAGGAATT	191
Db	21	SerLeuProValGlnIleThrValProGluLysIleArgSerIleIleLysGluGlyLe	40
Qy	192	GAATCGCAGGATCCTACAAAATTTGTAATTCAGAGGAAACCATATCTACTGTGAATTTAATG	251
Db	41	GluSerGlnAlaSerTyLysIleValIleGluGlyLysProTyThrValAsnLeuMet	60
Qy	252	CAAAAAAATTTTACCCCAATTTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTTAC	311
Db	61	GlnLysAsnPheLeuProHisAsnPheArgValTySerTySerGlyThrGlyLeMet	80
Qy	312	AAACCATTTGACCAAGATTTTACAGATTTTACAGATTTTACAGATTTTACAGATTTTAC	371
Db	81	LysProLeuAspGlnAspPheGlnAsnPheCysHisTyThrGlnGlyTyIleGluGlyTy	100
Qy	372	CCAAAATCTGGTGTGGTGTAGACATGTACTGACTCAGGGCGGTACTACAGTTTGA	431
Db	101	ProLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGlu	120
Qy	432	AATGTTAGTTATGAATAAGAACCCCTGGAGTCTTTCAGTTGGCTTTGAACATGTAAATTTAC	491
Db	121	AsnValSerTyGlyIleGluProLeuGluSerValGlyPheGluHisValIleTy	140
Qy	492	CAAGTAAACATAGAAGACAGATGTTTCCCTTATATATATATATATATATATATATATAT	551
Db	141	GlnValLysHisLysLysAlaAspValSerLeuTyAsnGluLysAspIleGluSerArg	160
Qy	552	GATCTGCTCTTAAATTTACAAAGCGCAGCAGCAGCAAGATTTTGCAGATATATAGAA	611
Db	161	AspLeuSerPheLysLeuGlnSerAlaGluProGlnGlnAspPheAlaLysTyIleGlu	180
Qy	612	ATGCATGTTATGTTGAAAAACAATTTGATATATATATATATATATATATATATATATAT	671
Db	181	MethIleValIleValGluLysGlnLeuTyAsnHisMetGlySerAspThrThrValVal	200
Qy	672	GCTCAAAAGTTTCCAGTGTGATGGATGACGAATGCTATTTTGTTCATTTAATATT	731
Db	201	AlaGlnLysValPheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsnIle	220
Qy	732	ACAAATTTCTCTCTTCTTCTGAGCTTTGGATAGATGAAATTAATAATTTGCAACACTGGA	791
Db	221	ThrIleIleLeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThrThrGly	240

Qy	792	GAAGCTAATGAGTATTATACACATTTTAAAGATGAAACAACTATCTTATCTTGTATTACGT	851
Db	241	GluAlaAsnGluLeuLeuHisThrPheLeuArgTyrLysThrSerTyLeuValLeuArg	260
Qy	852	CCTCATGATGTGGCATTCTTTTACTTGTTCACAGAAAAGTCAATTTATCTTGTGTCAACC	911
Db	261	ProHisAspValAlaPheLeuLeuValTyArgGluLysSerAsnTyValGlyAlaThr	280
Qy	912	TTTCAGGGAAGATGTGTGATGCAAACTATGAGGAGGTGTGTCTCTGACACCCCAAGACC	971
Db	281	PheGlnGlyMetCysAspAlaSerTyAlaGlyValValLeuHisProArgThr	300
Qy	972	ATAAGTCTGGAATCACTTCAGTTATTTAGTCTCAATTTAGCTTATGAGCTTATGTTGGGATC	1031
Db	301	IleSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGlyIle	320
Qy	1032	ACTTATGATGACATTTAAACAATGCCAGTCTCAGGAGCTCTGCAATTTATGATCCAGAA	1091
Db	321	ThrTyAspAspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnProGlu	340
Qy	1092	GCAATTCATTTCAGTGTGTGAAGATCTTTTAGTAACCTGCAGCTTCGAAGACTTTGCACAT	1151
Db	341	AlaIleHisPheSerGlyValLysIlePheSerAsnCysSerPheGluAspPheAlaHis	360
Qy	1152	TTTATTTCAAAGCAGAAAGTCCAGTCTCTTCAATTCAGCTCGCTTATGATCTTTTTC	1211
Db	361	PheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhePhe	380
Qy	1212	AAACACACAGCAGTGTGTGTAATGCAAGCTGGAAGCAGGAGGAGGTGTGCTGTGG	1271
Db	381	LysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCysGly	400
Qy	1272	ACTGAACAGGATGTGCCCTTTATGGAGAAACATGCTGTGATATTGCCCATGTAGATT	1331
Db	401	ThrGluGlnAspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPhe	420
Qy	1332	AAAGCCGGTTCAAACCTGTGCTGAAGACCAATGCTGTGATATTGCCCATGTAGATT	1391
Db	421	LysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLys	440
Qy	1392	GAAGATGTGTAGGCTTCTTGAAGATGCGACTCCCTCAATTTATGCAATGATCA	1451
Db	441	GluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyCysAsnGlySer	460
Qy	1452	TCGTGCATCATGCCAGAAAACCACTATGTTTCAGACTGGGCATCCGTGTGGAATCAA	1511
Db	461	SerAlaSerCysProGluAsnHisTyValGlnThrGlyHisProCysGlyLeuAsnGln	480
Qy	1512	TGGATCTGTATAGATGGATTTGTATGAGTGGGATAAACAATGATACAGACATTTGGC	1571
Db	481	TrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGly	500
Qy	1572	AAAGATGATGAGTTGGCTTCCAGATGTTTATTTTACCTTAATTTCAAAGACTGATGA	1631
Db	501	LysGluValGluPheGlyProSerGluCysTySerHisLeuAsnSerLysThrAspVal	520
Qy	1632	TCGTGAAACTGTGTATAGTATTAGATACACAGTGTGAAGTGTGAATCTTCAG	1691
Db	521	SerGlyAsnCysGlyIleSerAspSerGlyTyThrGlnCysGluAlaAspAsnLeuGln	540
Qy	1692	TGCGAAAAATTAATGTAATATGAGTAAATTTTATTTTATTTTATTTTATTTTATTTT	1751
Db	541	CysGlyLysLeuIleCysLysTyValGlyLysPheLeuLeuGlnIleProArgAlaThr	560
Qy	1752	ATTATTTATGCCAACATAAGTGCACATCTCTGATTTCTGTGGAATTTGTCAGTGTAT	1811
Db	561	IleIleTyAlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHis	580
Qy	1812	GCAGACACCAAGATGTGGATAAAGATGGAATCTTCTTGTGTGTGTTCAATTAAGTTTC	1871
Db	581	AlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCys	600
Qy	1872	AGGATCAAAGATGTGTGAGTCTTCTTACTACTTGGTTATGATTTGTTACTTACTGACAAATGC	1931

||||| ArgAsnGlnArgCysValSerSerSerTyrLeuGlyTyrAspCysThrThrAspLysCys 620  
Db  
QY 1932 AATGATAGAGTGATGCAATAACAAAGACACTGTCTACCTGTAGTGTCTTCATATTTACCT 1991  
|||||  
Db 621 AsnAspArgGlyValCysAsnAsnLysHisCysHisCysSerAlaSerTyrLeuPro 640  
QY 1992 CCAGATTGCTCAGTTCAATCAGATCTATGCGCTGGAGTATTGACAGTGGCAATTTT 2051  
Db 641 ProAspCysSerValGlnSerAspLeuTrpProGlySerIleAspSerGlyAsnPhe 660  
QY 2052 CCACCTGTAGCTATACAGCCAGACTCCCTGAAGGCGCTACATTGAGAACATTTACCAT 2111  
Db 661 ProProValAlaIleProAlaArgLeuProGluArgGlyIleGluAsnIleTyrHis 680  
QY 2112 TCCAAACCAATGAGATGGCCATTTTCTTATTTCATTCCTCTTCTTATTATTCTGTGTA 2171  
Db 681 SerLysProMetArgTrpProPhePheLeuPheIleProPhePheIlePheCysVal 700  
QY 2172 CTGATTGCTATAATGTTGAAGTTAAATTCCAAAGGAAAAAATGGAGAACTGAGGACTAT 2231  
Db 701 LeuIleAlaIleMetValLysValAsnPheGlnArgLysLysTrpArgThrGluAspTyr 720  
QY 2232 TCAAGCGATGAGCAACCTGAAAGTGAGAGTCAACCTAAAGG 2273  
Db 721 SerSerAspGluGlnProGluSerGluSerGluProLysGly 734  
RESULT 2  
ABJ19246  
ID ABJ19246 standard; protein; 734 AA.  
XX AC ABJ19246;  
XX DT 28-MAR-2003 (first entry)  
XX DE Human cancer/testis antigen - SEQ ID NO 19.  
XX KW Human; gene therapy; vaccine; cancer; cancer/testis antigen; CT antigen.  
XX OS Homo sapiens.  
XX PN WO200278526-A2.  
XX PD 10-OCT-2002.  
XX PF 29-MAR-2002; 2002WO-US009808.  
XX PR 30-MAR-2001; 2001US-0280718P.  
XX PR 20-APR-2001; 2001US-0285154P.  
XX PR 05-OCT-2001; 2001US-0327432P.  
XX PR 22-JAN-2002; 2002US-00054683.  
XX (LUDW-) LUDWIG INST CANCER RES.  
XX PA (CORR) CORNELL RES FOUND INC.  
XX PI Old LJ, Scanlan MJ, Chen Y;  
XX WPI: 2003-040608/03.  
XX N-PSDB; ABT15727.  
XX Diagnosing cancer comprises contacting a biological sample isolated from  
PT a subject with an agent that specifically binds to a nucleic acid  
PT molecule, its expression product or fragment or an antibody that binds to  
PT the product or fragment.  
XX Claim 36; Page 124-126; 155pp; English.  
XX The invention comprises a method for diagnosing cancer, the method  
XX involves detecting the DNA or protein sequences of human cancer/testis  
XX (CT) antigens that are disclosed in the invention. The method of the  
XX invention is useful for detecting/diagnosing, treating and monitoring a  
XX cancer or condition characterised by the expression of a human CT  
XX antigen. The present amino acid sequence represents a human CT antigen of

CC the invention  
SQ Sequence 734 AA;  
Alignment Scores:  
Pred. No.: 0 Length: 734  
Score: 3984.00 Matches: 734  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 83.07% Indels: 0  
DB: 6 Gaps: 0  
US-10-054-683-18 (1-2640) x ABJ19246 (1-734)  
QY 72 ATGTGGTCTTGTTCCTTCGCTCAGGGGCTCGGGGCTCGGATGGACAGTAATTTTGAT 131  
Db 1 MetTrpValLeuPheLeuLeuSerGlyLeuGlyLeuArgMetAspSerAsnPheAsp 20  
QY 132 AGTTTACCTGTGCNAATTACAGTTCGGAGAAATACGGTCAATATAAAGGAGGAATT 191  
Db 21 SerLeuProValGlnIleThrValProGluLysIleArgSerIleIleLysGluGlyLe 40  
QY 192 GAATCGCAGGCATCTACAAAATTGTAATTGAAGGGAACCATATACCTGTGAATTTAATG 251  
Db 41 GluSerGlnAlaSerTyrLysIleValIleGluGlyLysProTyrThrValAsnLeuMet 60  
QY 252 CAAAAAACTTTTACCCCATATTTTACAGTTTACAGTTTATAGTGCACAGGAATTTATG 311  
Db 61 GlnLysAsnPheLeuProHisAsnPheArgValTyrSerTyrSerGlyThrGlyLeuMet 80  
QY 312 AAACCACTTGACCAAGATTTTTCAGAAATTTCTGCCACTTACCAGGGTATATTGAAGTTAT 371  
Db 81 LysProLeuAspGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGlyTyr 100  
QY 372 CCAAAATCTGT 431  
Db 101 ProLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGlu 120  
QY 432 AATGTTAGTTATGGAATAGAACCCCTGGAGTCTTTCAGTTCGCTTTGAACATGTAAATTTAC 491  
Db 121 AsnValSerTyrGlyIleGluProLeuGluSerSerValGlyPheGluHisValIleTyr 140  
QY 492 CAAGTAAACATAAGAAAGCAGATGTTTCCTTATATATATAGAGGATTTTGAATCAAGA 551  
Db 141 GlnValLysHisLysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSerArg 160  
QY 552 GATCTGTCCTTTTAAATTTACAAAGCGCAGAGCCAGCAAGATTTTGCAAAGTATATAGAA 611  
Db 161 AspLeuSerPheLysLeuGlnSerAlaGluProGlnGlnAspPheAlaLysTyrIleGlu 180  
QY 612 ATGCATGTTATAGTTGAAAAACAATTTGTATAATCATATGGGTCTGTATACAACTGTTGTC 671  
Db 181 MetHisValIleValGluLysGlnLeuTyrAsnHisMetGlySerAspThrThrValVal 200  
QY 672 GCTCAAAAAGTTTTCAGTTGATTGAGCAATGCTATTTTGTTCATTATTTTAAATTT 731  
Db 201 AlaGlnLysValPheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsnIle 220  
QY 732 ACAATTTATCTGCTTCCTTCATTTGGAGCTTTGGATGAAATAAATTTGCAACCACTGGA 791  
Db 221 ThrIleLeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThrThrGly 240  
QY 792 GAAGCTAATCAGTTATTACACACATTTTTTAAGATGGAACACATCTTATCTTTGTTTACGT 851  
Db 241 GluAlaAsnGluLeuLeuHisThrPheLeuArgTrpLysThrSerTyrLeuValLeuArg 260  
QY 852 CTTTCATGATGTGGCATTTTTCCTTTTACAGAAAAAGCAAAATTTATTTGTTGGTCAACC 911  
Db 261 ProHisAspValAlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAlaThr 280  
QY 912 TTTTCAAGGAGATGTTGTGATGCAAACTATGCAGAGGTGTTGTTCTGCCCCCGACGACC 971  
Db 281 PheGlnGlyLysMetCysAspAlaAsnTyrAlaGlyGlyValValLeuHisProArgThr 300

Qy 972 ATAACTCTGGAATCACCTTGCAGTTATTTAGCTCAATATTAGCCTTAGTATGGGATC 1031  
Db |||||  
Qy 301 ILeSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGlyIle 320  
Db |||||  
Qy 1032 ACTTATGATGACATTAAACATGCCAGTCTCAGAGCTGTCTGCATTATGAATCCAGAA 1091  
Db |||||  
Qy 321 ThrTyrAspAspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnProGlu 340  
Db |||||  
Qy 1092 GCAATTCATTTCACTGGTGTCAAGATCTTTAGTAACATGCAGCTTCGAGACCTTGGCACAT 1151  
Db |||||  
Qy 341 AlalleHisPheSerGlyValLysIlePheSerAsnCysSerPheGluAspPheAlaHis 360  
Db |||||  
Qy 1152 TTTATTTCAACAGCAGAGTCCAGTGTCTTACCAATCAGCTCGCTTAGATCTCTTTTTC 1211  
Db |||||  
Qy 361 PheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspPhePhe 380  
Db |||||  
Qy 1212 AAACAGCAGCAGTGTGTGTAATGCAAGCTGGAAGCAGAGAGGTGTGATGTGGG 1271  
Db |||||  
Qy 381 LysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCysGly 400  
Db |||||  
Qy 1272 ACTGAACAGGATTCGCCCTTATTGGAGAACATCTCTGTGATATTGCCACATGTAGATT 1331  
Db |||||  
Qy 401 ThrGluGlnAspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPhe 420  
Db |||||  
Qy 1332 AAAGCCGCTTCAAACTGTCTGCTGAAGGACCATGCTGCGAAAACTGTCTATTATGTCAAAA 1391  
Db |||||  
Qy 421 LysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLys 440  
Db |||||  
Qy 1392 GAAAGATGTGTAGCCCTTCTTTGAAGAATGCGACCTCCCTGGAATATTGCAATGGATCA 1451  
Db |||||  
Qy 441 GluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGlySer 460  
Db |||||  
Qy 1452 TCTGCATCATGCCAGAAACCACTATGTTGAGTGGCATCGGTGGTGGACTGAATCAA 1511  
Db |||||  
Qy 461 SerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGln 480  
Db |||||  
Qy 1512 TGGATCTGTATAGATGGAGTTTGTATAGTGGGATAAACAATGTACAGACACATTTGGC 1571  
Db |||||  
Qy 481 TrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGly 500  
Db |||||  
Qy 1572 AAAGAAGTAGAGTTGGCCCTTCAGATGTTATTCTCACTTAATTCAAAGACTGATGTA 1631  
Db |||||  
Qy 501 LysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspVal 520  
Db |||||  
Qy 1632 TCTCGAACTGTGTATTAAGTATTTCAGGATACACAGCTGTGAGCTGCAATCTGCAG 1691  
Db |||||  
Qy 521 SerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGln 540  
Db |||||  
Qy 1692 TGCGGAAAAATTAATATGTAATATGTAAGTAAATTTTATTACAAATTTCCAAAGAGCCACT 1751  
Db |||||  
Qy 541 CysGlyLysLeuIleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAlaThr 560  
Db |||||  
Qy 1752 ATTTATTTATGCAACATAAGTGGACATCTTCGCAATGTGTGGAATTTGCCAGTGCAT 1811  
Db |||||  
Qy 561 IleIleTyrAlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHis 580  
Db |||||  
Qy 1812 GCAGACGCCAAAGATCTGATAAAGATGGAATCTTCTGCTTCAATTAAGTTGTC 1871  
Db |||||  
Qy 581 AlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCys 600  
Db |||||  
Qy 1872 AGGAATCAAGATGTGTGAGTTCTTCACTATGTTGGTTATGATTGTCTACTGCAAAATGC 1931  
Db |||||  
Qy 601 ArgAsnGlnArgCysValSerSerTyrLeuGlyTyrAspCysThrThrAspLysCys 620  
Db |||||  
Qy 1932 AATGATAGGTGTATGCAATAACAAAGACATGTCTAGTGTCTCATATTTACCT 1991  
Db |||||  
Qy 621 AsnAspArgGlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeuPro 640  
Db |||||  
Qy 1992 CCAGATTGCTCAGTTCAATCAGATCTATGGCTGTGGAGTATTGACAGTGGCAATTT 2051  
Db |||||  
Qy 641 ProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsnPhe 660  
Db |||||

Qy 2052 CCACCTCTAGCTATPACAGCCAGACTCCCTGAAAGCGCTACATTGAGAACATTTACCAT 2111  
Db |||||  
Qy 661 ProProValAlaIleProAlaArgLeuProGluArgTyrIleGluAsnIleTyrHis 680  
Db |||||  
Qy 2112 TCNAACCAATGAGATGCCATTTTCTTATTATTCATTCCTTCTTATTTATTTCTGTGTA 2171  
Db |||||  
Qy 681 SerLysProMetArgTrpProPhePheLeuPheIleProPhePheIlePheCysVal 700  
Db |||||  
Qy 2172 CTGATTCCTTAATATGGTGAAGTTAATTTCCAAAGGAAAAAATCGAGAACTGAGGACTAT 2231  
Db |||||  
Qy 701 LeuIleAlaIleMetValLysValAsnPheGlnArgLysLysTrpArgThrGluAspTyr 720  
Db |||||  
Qy 2232 TCAAGCGATGAGCAACTGAAAGTGAGAGTCAACTAAAGGG 2273  
Db |||||  
Qy 721 SerSerAspGluGlnProGluSerGluSerGluProLysGly 734  
Db |||||

## RESULT 3

ADB75186  
ID ADB75186 standard; protein; 734 AA.  
XX ADB75186;  
AC ADB75186;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Prostate cancer marker protein.  
XX  
KW Prostate; cancer; cytostatic; gene therapy; marker.  
XX Homo sapiens.  
XX  
FN WO2003009814-A2.  
XX  
PD 06-FEB-2003.

25-JUL-2002; 2002WO-US023913.

25-JUL-2001; 2001US-0307982P.

22-AUG-2001; 2001US-0314356P.

25-SEP-2001; 2001US-0325020P.

12-DEC-2001; 2001US-0341746P.

05-MAR-2002; 2002US-0362158P.

(MILL-) MILLENNIUM PHARM INC.

Schlegel R, Monahan JB, Endege WO, Gannavarapu M, Gorbacheva B;

Hoersh S, Kamatkar S, Womsey AM, Glatt K, Zhao X, Anderson D;

WPI; 2003-248033/24.

New nucleic acid molecule, useful for diagnosing or treating prostate

cancer.

Disclosure; SEQ ID NO 10; 99pp; English.

The invention relates to newly discovered cancer markers associated with

the cancerous state of prostate cells. Also disclosed is a method of

assessing whether a patient is afflicted with prostate cancer. The method

of the invention involves assessing whether a patient is afflicted with

prostate cancer by comparing the level of expression of a marker in a

patient sample and the normal level of expression of the marker in a

control non-prostate cancer sample, where a significant increase in the

level of expression of the marker in the patient sample and the normal

level indicates that the patient is afflicted with prostate cancer.

Nucleic acids of the invention are useful for diagnosing or treating

prostate cancer, and may be useful in gene therapy. Sequences given in

CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence

data for this patent did not form part of the printed specification, but

was obtained in electronic format directly from WIPO at

ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 734 AA;

Alignment Scores:



Pred. No.:	0	Length:	734
Score:	3984.00	Matches:	734
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	83.07%	Indels:	0
DB:	7	Gaps:	0

US-10-054-683-18 (1-2640) x ADB75186 (1-734)

QY	72	ATGTGGGTCTGTTCTCTGCTCAGCGGCTCGCGGGCTCGCGATGAGCAGTAATTTTGAT	131
DB	1	MetTrpValLeuPheLeuLeuSerGlyLeuGlyLeuArgMetAspSerAsnPheAsp	20
QY	132	AGTTTACCTGTGCAAAATTTACAGTTCCGGAGAAAATACGGTCAATTAATAAGGAAGGAATT	191
DB	21	SerLeuProValGlnIleThrValProGluLysIleArgSerIleIleLysGluGlyIle	40
QY	192	GAATCGCAGGCATCTCAAAATTTGTAATGAAGGGAACCATATACTGTGTAATTAATG	251
DB	41	GluSerGlnAlaSerTyrLysIleValIleGluGlyLysProTyrThrValAsnLeuMet	60
QY	252	CAAAAAAATTTTACCCCAATAATTTAGAGTTTACAGTTATAGTGGCAGCAGGAATTAATG	311
DB	61	GlnLysAsnPheLeuProHisAsnPheArgValTyrSerTyrSerGlyThrGlyIleMet	80
QY	312	AAACCACTTGACCAAGATTTTCAGAAATTTCTGCCACTTACCAGGGGTATATTGAAGGTTAT	371
DB	81	LysProLeuAspGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGlyTyr	100
QY	372	CCAAATCTGTGGTGATGTTAGCACATGTACTGGACTCAGGGCGGTACTACAGTTGAA	431
DB	101	ProLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGlu	120
QY	432	AATGTTAGTTATGAATAGAACCCCTCGAGTCCTTCAGTTGGCTTTGAACATGTAAATTTAC	491
DB	121	AsnValSerTyrGlyIleGluProLeuGluSerValGlyPheGluHisValIleTyr	140
QY	492	CAAGTAAACATAAGAAGCAGATGTTTCTCTTATATATGAAGAGGATATTGAATCAAGA	551
DB	141	GlnValLysHisLysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSerArg	160
QY	552	GATCTGTCCTTTAAATTTACAAAGCGCAGAGCCACAGCAAGATTTTGCAGAGTATAGAA	611
DB	161	AspLeuSerPheLysLeuGlnSerAlaGluProGlnAspPheAlaLysTyrIleGlu	180
QY	612	ATGCATGTTATAGTTGAAAACAATGTATATCATATGGGCTCTGATACAACTGTTGTC	671
DB	181	MethHisValIleValGluLysGlnLeuTyrAsnHisMetGlySerAspThrThrValVal	200
QY	672	GCTCAAAAAGTTTCCAGTTGATGGAATGCGAATGCTATTTTGTTCATTTTAATATT	731
DB	201	AlaGlnLysValPheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsnIle	220
QY	732	ACAATTAATCTGCTTCAATGGACCTTGGATAGATGAATAAATAATTTGCACACCTGGA	791
DB	221	ThrIleIleLeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThrThrGly	240
QY	792	GAAGCTAATGAGTTATTACACACATTTTAAAGATGGAACATCTTATCTTTGTTTACGT	851
DB	241	GluAlaAsnGluLeuLeuHisThrPheLeuArgTrpLysThrSerTyrLeuValLeuArg	260
QY	852	CCTCATGATGGCAATTTTACTTGTGTTTACAGAAAAGTCAAAATTAATGTTGGTGCACAC	911
DB	261	ProHisAspValAlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAlaThr	280
QY	912	TTTCAAGGAAGATGTGTGATGCAAACTATGCCAGAGGTGTGTTCTGCAACCCAGAAC	971
DB	281	PheGlnGlyLysMetCysAspAlaAsnTyrAlaGlyGlyValValLeuHisProArgThr	300
QY	972	ATAAGTCTGGAATCACCTGCAAGTTATTTAGCTCAATTAATGAGCCCTAGTATGGGATC	1031
DB	301	IleSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGlyIle	320

QY	1032	ACTTATGATGACATTAACAAAATGCCAGTGTCTCAGGAGCTGTCTGCATTTATGAATCCAGAA	1091
DB	321	ThrTyrAspAspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnProGlu	340
QY	1092	GCAATTCATTTTCAGTGTGTGAAGATCTTTAGTAACATGCAGCTTCGAGACTTTTGCACAT	1151
DB	341	AlaIleHisPheSerGlyValLysIlePheSerAsnCysSerPheGluLeuPheAlaHis	360
QY	1152	TTTATTTTCAAGCAGAAAGTCCCAAGTGTCTTTCACAATCAGCCTCGCTTAGATTCCTTTTTC	1211
DB	361	PheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuLeuAspProPhePhe	380
QY	1212	AAACAGCAACAGTGTGTGTGTAATGCAAAAGCTGGAACAGAGAGAGTGTGACTGTGGG	1271
DB	381	LysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCysGly	400
QY	1272	ACTGAACAGAGATGTGCCCTTTATTGGAGAAACATGCTGTGATATTGCCACATGTAGATTT	1331
DB	401	ThrGluGlnAspCysAlaLeuIleGlyLysThrCysCysAspIleAlaThrCysArgPhe	420
QY	1332	AAAGCCGTTTCAAACTGTGCTGAAGGACCATGTCTGCGAAAACCTGTCTATTATTATCTCAA	1391
DB	421	LysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLys	440
QY	1392	GAAGAATGTGTAGGCCCTTCTTTGAAGATGCGACCTCCCTGAAATATTGCAATGGATCA	1451
DB	441	GluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGlySer	460
QY	1452	TCGTCATCATGCCAGAAAACCACTATGTCAGACTCGGCATCCGTGGACTCAATCAAA	1511
DB	461	SerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGln	480
QY	1512	TGGATCTGTATAGATGAGTGTGTATGAGTGGGATAAACAATGTACAGACACATTTGGC	1571
DB	481	TrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGly	500
QY	1572	AAAGAATAGATTTGGCCCTTCAGAAATGTTATTCTCACCTTAATTCAGAGCTGATGTA	1631
DB	501	LysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspVal	520
QY	1632	TCGTGAACCTGTGTATAAGTATTCAGGATACACAGAGTGTGAAGCTGCAATCTGCAG	1691
DB	521	SerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGln	540
QY	1692	TGCGAAAAATTAATATGTAATATGTAAGTAAATTTTATTACAAATTTCCAGAGCCACT	1751
DB	541	CysGlyLysLeuIleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAlaThr	560
QY	1752	ATTATTTATGCCAACATAAGTGGACATCTCTGCAATTCCTGTGGAATTTGCCAGTATCAT	1811
DB	561	IleIleTyrAlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHis	580
QY	1812	GCAGACACCAAGATGTGTAATAAGATGGAACCTTCTGTGGTTCAATAAGGTTTGC	1871
DB	581	AlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCys	600
QY	1872	AGCAATCAAAAGATGTGTGAGTCTTCTCATCTTGGTTATGATTTGATCTACTGACAAATGC	1931
DB	601	ArgAsnGlnArgCysValSerSerTyrLeuGlyTyrAspCysThrThrAspLysCys	620
QY	1932	AATGATAGAGGTGATGCCAATAACAAAGACCTGTCTGCTAGTGTCTTATATTTACCT	1991
DB	621	AsnAspArgGlyValCysAsnAsnLysHisCysHisCysSerAlaSerTyrLeuPro	640
QY	1992	CCAGATTTGCTCAGTTCAATCAGATCTATGCCCTGTGGGAGTATTGACAGTGGCAATTTT	2051
DB	641	ProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsnPhe	660
QY	2052	CCACTGTAGCTATACAGCCAGACTCCCTGAAAGCGCTACATTGAGAACATTTACCAT	2111
DB	661	ProProValAlaIleProAlaArgLeuProGluArgTyrIleGluAsnIleTyrHis	680
QY	2112	TCCAACCAATGAGATGGCCATTTTCTTATTATTCATTCCTTTCTTTTATTATTTCTGTGTA	2171



```
Db 341 GluAlaIleHisPheSerGlyValLysIlePheSerAsnCysSerPheGluAspPheAla 360
QY 1149 CATTATTATTTCAAGCAGAGTCCAGTGTCTTCACAATCAGCTCGCTTAGATCTCTTTT 1208
Db 361 HisPheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhe 380
QY 1209 TTCAAAACAGCAGTGTGTGTAATGCAAGCTGGAAGCAGGAGGAGTGTGACTGT 1268
Db 381 PheLysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCys 400
QY 1269 GGCACCTGAACAGATTGTGCGCTTATTGGAAACATGCTGTGATATTGCCACATGTAGA 1328
Db 401 GlyThrGluGlnAspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArg 420
QY 1329 TTTAAAGCCGTTCAAACTGTGTGAAGGACCATGTGCGAAACCTGTCTATTATTGTCA 1388
Db 421 PheLysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSer 440
QY 1389 AAAGAAAGATGTGTAGGCCCTTCCTTTGAAGAAATGCGACCTCCCTGAATATTGCAATGGA 1448
Db 441 LysGluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGly 460
QY 1449 TCATCTGCATCATGCCAGAAACCACTATGTCAGACTGGGCATCGGTGTGACTGAAT 1508
Db 461 SerSerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsn 480
QY 1509 CAATGATCTGTATAGATGAGTTGTATGAGTGGGATAAACAATGTCACAGACACATTT 1568
Db 481 GlnTrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPhe 500
QY 1569 GCGAAAGAAAGTAGAGTTGGCCCTTCAGAAATGTTATTCTCACCTTAATTCAAAGACTGAT 1628
Db 501 GlyLysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAsp 520
QY 1629 GTATCTGGAAACTGTGTATAGTGATTAGGATACACAGTGTGAGCTGACATCTGT 1688
Db 521 ValSerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeu 540
QY 1689 CAGTGGGAAATTAATATGTAATGTAAGTAAATTTTATTACAAATTCACAGGCC 1748
Db 541 GlnCysGlyLysLeuIleCysIleTyrValGlyLysPheLeuLeuGlnIleProArgAla 560
QY 1749 ACTATTATTTATGCCAACAATAGTGACATCTCTGCATTTCTGTGGAATTTGCCAGTGAT 1808
Db 561 ThrIleIleTyrAlaIleHisSerGlyHisLeuCysIleAlaValGluPheAlaSerAsp 580
QY 1809 CATGCGACAGCCAAAGATGTGGATAAAGATGGAACCTCTTGTGTGTTTCAAAATAGGTT 1868
Db 581 HisAlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysVal 600
QY 1869 TGCAGGAATCAAGATGTGTGAGTCTTCTCATCTTGGTTATCATTTGCTACTACTGCAAA 1928
Db 601 CysArgAsnGlnArgCysValSerSerSerTyrLeuGlyTyrAspCysThrThrAspLys 620
QY 1929 TGCAAATGATAGGTGTATGCAATAACAAAGCACTGTCACTGTAGTGTCTCATATTTA 1988
Db 621 CysAsnAspArgGlyValCysAsnAsnLysHisCysHisCysSerAlaSerTyrLeu 640
QY 1989 CTCTCCAGATTCTCAGTTCAATCAGATCTATAGCCCTGGTGGAGTATTTGACATGGCAAT 2048
Db 641 ProProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsn 660
QY 2049 TTTCCACCTGTAGTATACAGCCAGACTCCCTGGAAGCGCTACATTGGAGACATTTAC 2108
Db 661 PheProProValAlaIleProAlaArgLeuProGluArgTyrIleGluAsnIleTyr 680
QY 2109 CATTCCAAACCAATGAGATGGCAATTTTCTTATTCAATCTCTTCTTATTATTCTGT 2168
Db 681 HisSerLysProMetAlaGlyTrpProPhePheLeuPheIleProPhePheIlePheCys 700
QY 2169 GTACTGATTCCTATAATGGTGAAGTTAATTTCCAAAGGAAAAAATGGAGAACTGAGGAC 2228
```

```
Db 701 ValLeuIleAlaIleMetValLysValAsnPheGlnArgLysLysTrpArgThrGluAsp 720
QY 2229 TATTCAAGCGATGAGCAACCTGAAAGTGTGAGAGTGAACCTAAAGGG 2273
Db 721 TyrSerSerAspGluGlnProGluSerGluSerGluProLysGly 735
```

## RESULT 5

```
ADB75182
ID ADB75182 standard; protein; 753 AA.
XX ADB75182;
AC ADB75182;
XX 04-DEC-2003 (first entry)
XX Prostate cancer marker protein.
XX Prostate; cancer; cytostatic; gene therapy; marker.
XX Homo sapiens.
XX WO2003009814-A2.
XX 06-FEB-2003.
XX 25-JUL-2002; 2002WO-US023913.
XX 25-JUL-2001; 2001US-0307982P.
XX 22-AUG-2001; 2001US-0314356P.
XX 25-SEP-2001; 2001US-0325020P.
XX 12-DEC-2001; 2001US-0341746P.
XX 05-MAR-2002; 2002US-0362158P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
XX Hoerhn S, Kamatkar S, Wonsey AM, Giatt K, Zhao X, Anderson D;
XX WPI; 2003-248033/24.
XX New nucleic acid molecule, useful for diagnosing or treating prostate
XX cancer.
XX Claim 4; SEQ ID NO 6; 99pp; English.
```

The invention relates to newly discovered cancer markers associated with the cancerous state of prostate cells. Also disclosed is a method of assessing whether a patient is afflicted with prostate cancer. The method of the invention involves assessing whether a patient is afflicted with prostate cancer by comparing the level of expression of a marker in a patient sample and the normal level of expression of the marker in a control non-prostate cancer sample, where a significant increase in the level of expression of the marker in the patient sample and the normal level indicates that the patient is afflicted with prostate cancer. Nucleic acids of the invention are useful for diagnosing or treating prostate cancer, and may be useful in gene therapy. Sequences given in CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 753 AA;

Alignment Scores:			
Pred. No.:	0	Length:	753
Score:	3900.00	Matches:	720
Percent Similarity:	97.44%	Conservative:	4
Best Local Similarity:	96.90%	Mismatches:	13
Query Match:	81.32%	Indels:	6
DB:	7	Gaps:	1

US-10-054-683-18 (1-2640) x ADB75182 (1-753)

QY 45 TCCGGCTGGGACCCAGGACTTCAAGCCATGTGGGTCTTGTCTCTCAGCGGCTCGGC 104











polypeptides are used to identify modulators. A peptide, phosphopeptide, small organic molecule or an antibody is used to inhibit aberrant activity of a 56294 or 56293-expressing cell, located in a cancerous or precancerous tissue. A compound that modulates the activity or expression of the polynucleotides are used to treat or prevent a disorder characterized by aberrant activity of a 56294 or 56293-expressing cell, in a subject. The present sequence represents the human 56294 protein.

Sequence 820 AA;

Alignment Scores:	
Pred. No.:	1.0e-127
Score:	1530.00
Percent Similarity:	57.84%
Best Local Similarity:	40.87%
Query Match:	31.90%
DB:	6
Length:	820
Matches:	318
Conservative:	132
Mismatches:	266
Indels:	62
Gaps:	20

US-10-054-683-18 (1-2640) x ABR39696 (1-820)

QY	6	CGCACTTCCAAC	TGCCCTGT	TAACCA	CCAACTGC	CCCTATT	TCGCGCTGGG	ACCAGG	ACTT	65
DB	22	ArgThr	ArgGly	Cys	-----	-----	-----	-----	-----	71
QY	66	CAAGCCATG	TGGGTCT	TTGTT	---CTG	CTACG	GGGCTCG	CGGCGT	CGGATG	119
DB	32	ArgThr	MetPhe	ArgLeu	TrpLeu	LeuLeu	AlaGly	Leu	CysGly	51
QY	120	AGTAATTTG	ATGATTAC	CTTCCG	GAGAAAT	ACGGTCA	CAATAA	TAATA	TA	179
DB	52	ProGly	PheGln	AsnSer	LeuLeu	GlnIle	ValIle	ProGlu	LysIle	71
QY	180	AGGAAAGAA	TTGAA	-----	-----	TCG	CAGCAT	CTCTACA	AAATTTGA	230
DB	72	AsnAsp	SerSer	GluIle	GlyTrp	GluGln	IleSer	TyrIle	IlePro	91
QY	231	CCATATAC	TGTGAAT	TTAAT	CCAAAAC	CTTTTAC	CCCATAA	TATTTAG	AGTTTAC	290
DB	92	LeuTyr	ThrVal	HisLeu	GlyGln	ArgTyr	PheLeu	AlaAsp	AsnPhe	111
QY	291	TATAGTGC	ACAGGA	TTATGAA	ACCATTG	ACCAAGAT	TTTCAGAA	TTTTCG	CACTAC	350
DB	112	TyrAsn	--Gln	GlySer	MetAsn	ThrTyr	SerSer	AspIle	GlnThr	130
QY	351	CAAGGGTAT	TTAGGTT	ATCCAAAT	CTGTG	TGTATG	TTAGC	ACATG	TACTG	410
DB	131	GlnGly	AsnIle	GluGly	TyrPro	AspSer	MetVal	ThrLeu	SerThr	150
QY	411	AGGGCGT	PACTAC	AGTTG	AAATGT	TGATTG	GAATAGA	ACCCCTG	GAGTCTC	470
DB	151	ArgGly	IleLeu	GlnPhe	GluAsn	ValSer	TyrGly	IleGlu	ProLeu	170
QY	471	GGCTTTG	AACTGTA	TTTACC	AGTAAA	CATAAG	AAAGCAG	ATGTTTCC	TATATAAT	530
DB	171	GluPhe	GlnHis	ValLeu	TyrLys	LeuLys	AsnGlu	AspAsn	AspIle	190
QY	531	GAGAAGG	ATATTGA	TGATCA	AG	-----	-----	GATCTG	CTCTTAA	581
DB	191	AspArg	SerLeu	LysGlu	GlnPro	MetAsp	AspAsn	IlePhe	IleSer	210
QY	582	CCACAGC	AAAGAT	-----	-----	TTT	CAAACTAT	AGAAATG	CATGTTATAG	635
DB	211	ProAla	ValPro	AspLeu	PhePro	LeuTyr	LeuGlu	MetHis	IleVal	230
QY	636	TTGTAT	ATCAT	TGCGGT	CTGTAT	CAACAT	GTGTTG	TCTCGCT	CAAAAAG	695
DB	231	LeuTyr	AspTyr	TrpGly	SerAsp	SerMet	IleVal	ThrAsn	LysVal	250
QY	696	GGATTG	ACGAAT	GCTAT	TTTGT	TTTCA	TTTAA	TATTA	CAATATCTG	755
DB	251	GlyLeu	AlaAsn	SerMet	PheThr	GlnPhe	LysVal	ThrIle	ValLeu	270
QY	756	CTTTGG	ATAGAT	GAATAA	ATAAT	TGC	ACCCTG	GAGAA	GCTTAAT	815

Db	271	LeuTrpSerAspGluAsnLysIleSerThrValGlyGluAlaAspGluLeuGlnLys	290
QY	816	TTTTAAAGATGAAACATCTTATCTGCTTTTACGTCCTCATGTGCGCATTTTACTT	875
Db	291	PheLeuGluTrpGlyGlnSerTyrLeuAsnLeuArgProHisAspIleAlaTyrLeuLeu	310
QY	876	GTTTCAGAGAAAAGTCAAAATTATGTGGTCAACCTTTCAAGGGAAGATGTGTGATGCA	935
Db	311	IleTyrMetAspTyrProArgTyrLeuGlyAlaValPheProGlyThrMetCysIleThr	330
QY	936	AACTATGAGGAGGTGTTCTGCACCCGAGAACCAATAAGTCTGGAACTCACTTGCAGTT	995
Db	331	ArgTyrSerAlaGlyValAlaLeuTyrProLysGluIleThrLeuGluAlaPheAlaVal	350
QY	996	ATTTAGCTCAATTATGAGCCTTAGTATGGGATCACTTATGATGACATTTAACAAAATGC	1055
Db	351	IleValThrGlnMetLeuAlaLeuSerLeuGlyIleSerTyrAspAspProLysLysCys	370
QY	1056	CAGTCTCAGGAGCTGCTGCAATATGAATCCAGAACCAATTCATTTCAGTGTGTGAAG	1115
Db	371	GlnCysSerGluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValLys	390
QY	1116	ATCTTTAGTAATCTGCAGCTTCCGAAGCTTTGTCACATTTATTTCAAAGCAGAAGTCCCAG	1175
Db	391	ThrPheSerSerCysSerLeuArgSerPheGlnAsnPheIleSerAsnValGlyValLys	410
QY	1176	TGCTTTCAATCAGCCTCGCTTAGATCTCTTTTCAAACAG---CAAGCAGTGTGTGGT	1232
Db	411	CysLeuGlnAsnAsnThrArg-----ProGlnLysLysSerProLysProValCysGly	428
QY	1233	AATGCAAGCTTGGAGCAGGAGAGAGTGTGACTGTGGGACTCAACAGGATTTGCGCCTT	1292
Db	429	AsnGlyArgLeuGluGlyAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly---	447
QY	1293	ATTGGAGAACATGCTGTGATTTGGCCATGTAGATTTTAAACCGGTTCAAACTGTGCT	1352
Db	448	---ProAlaSerCysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyr	466
QY	1353	GAAGGACCATGCTCGAAAACTGCTATTATTATGTCAAAGAAAGATGTGTAGCCCT---	1409
Db	467	LysGlyLeuCysCysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLys	486
QY	1410	TCCTTTGAAGATCGCACCTCCCTGTAATTTGCAATGGATCATCTGCATCATGCCAGAA	1469
Db	487	AlaHisProGluCysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyPro	506
QY	1470	AACCATCTGTTTCAGACTGGGCATCGTGTGGACTGAATCAATGATCTGTATAGATGA	1529
Db	507	AspIleThrLeuLeuAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGly	526
QY	1530	GTTTGTATGATGGGATAAACAATGTACAGACACATTTGGCAAGAAGTAGAGTTTGGC	1589
Db	527	AspCysHisAspLeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAla	546
QY	1590	CCITCAGATGTTTCTCACCTTAATCAAGACATGATGTATCTGGAAACTGTGGTATA	1649
Db	547	ProPheAlaCysTyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnGlyArg	566
QY	1650	AGT---GATTTCAGGATACACAGCTGTGAAGCTGCAACATCTCCAGTCGGAAAAATATA	1706
Db	567	AspArgAsnAsnLysTyrValPheCysGlyTyrArgAsnLeuIleCysGlyArgLeuVal	586
QY	1707	TGTAATATATGAGTAAATTTTTTATPACAAATTTCCAAGAGCCACTATTATTATGCCAAC	1765
Db	587	CysThrTyrProThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrAlaPhe	606
QY	1767	ATAAGTGGACATCTCTGCATCTGCTGGAATTTGGCAGTGTATCATGCAGACCCAAAAG	1825
Db	607	ValArgAspSerValCysIleThrValAspTyrLysLeuProArgThrValProAspPro	626
QY	1827	ATGTGGATAAAGATGGAACTTCTTGTGGTTCAAATAAGGTTTTCAGGAATCAAGATGT	1885





QY 420 CTACAGTTTCAAAATGTTAGTTATGGAATAGAACCCCTGGAGTCTTCAGTTGGCTTTGAA 479  
DB 121 LeuGlnPheGluAsnValSerTyrGlyLeuProLeuGluSerAlaValGluPheGln 140  
QY 480 CATGTAAATTACCAAGTAAACATATAAGAACAGAGATGTTTCCTTATATATATGAGAAGGAT 539  
DB 141 HisValLeuTyrLysLeuLysAsnGluAspAsnAspIleAlaIlePheIleAspArgSer 160  
QY 540 ATTGAATCAAGA-----GATCTGCTCTTTAAATTAACAAAGCCAGGACACAGCAA 590  
DB 161 LeuLysGluGlnProMetAspAsnIlePheIleSerGluLysSerGluProAlaVal 180  
QY 591 GAT-----TTTGCAAGATATAGAAATGCATGTTATAGTTGAAACAAATTTGATATAT 644  
DB 181 ProAspLeuPheProLeuTyrLeuGluMetHisIleValValAspLysThrLeuLysArg 200  
QY 645 CATATGGGTCTGATACAACTGTTGCGCTCAAAAAGTTTCCAGTTGATGGATGAGC 704  
DB 201 TyrTrpGlySerAspSerMetIleValThrAsnLysValIleGluIleValGlyLeuAla 220  
QY 705 AATGCTATTATTTGTTTCAATTAATATACAAATTTATCTGCTCTTCATTCAGTGGCTTTGGATA 764  
DB 221 AsnSerMetPheThrGlnPheLysValThrIleValLeuSerSerLeuGluLeuTrpSer 240  
QY 765 GATGAAATATAAATTCACCACTGAGAGAGCTAATGAGTTATTACACACATTTTAAAGA 824  
DB 241 AspGluAsnLysIleSerThrValGlyGluAlaAspGluLeuGlnLysPheLeuGlu 260  
QY 825 TGGAAACACTTATCTGTTTACGCTCTCATGATGGCGATTTTACTTGTGTTTACAGA 884  
DB 261 TrpLysGlnSerTyrLeuAsnLeuArgProHisAspIleAlaTyrLeuLeuIleTyrMet 280  
QY 885 GAAAGATCAAAATATGTTGCTGCAACCTTTCAAGGGAAGATGTGATGCAAACTATCCA 944  
DB 281 AspTyrProArgTyrLeuGlyAlaValPheProGlyThrMetCysIleThrArgTyrSer 300  
QY 945 GGAGTGTGTTGTCACCCCAAGAACCAATAGTCTGGAATCACTTGCAGTTATTTTACT 1004  
DB 301 AlaGlyValAlaLeuTyrProLysGluIleThrLeuGluAlaPheAlaValIleValThr 320  
QY 1005 CAATATTATTAGCTTAGTATGGGATCACTTATGATGACATTAACAATGCGAGTGTCA 1064  
DB 321 GlnMetLeuAlaLeuSerLeuGlyIleSerTyrAspAspProLysLysCysGlnCysSer 340  
QY 1065 GGAGCTGTCTGATTAATGATCAAGCAATTCATTTTCAGTGGTGTGAAGATCTTTAGT 1124  
DB 341 GluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValLysThrPheSer 360  
QY 1125 AACTGCAGCTTGAAGACTTTCACATTTTATTTCAAGCAGAGATCCAGTGTCTTCAAC 1184  
DB 361 SerCysSerLeuArgSerPheGlnAsnPheIleSerAsnValGlyValLysCysLeuGln 380  
QY 1185 AATCAGCTCGCTTAGATCTTTTTCACACAGCAGCAGTGTGTTGTAATGCAAGCTG 1244  
DB 381 AsnLysProGlnMetGlnLys-----LysSerProLysProValCysGlyAsnGlyArgLeu 399  
QY 1245 GAACAGAGAGAGGATGTGATGTGGGACTGAACAGAGATTTGTCCTTATTGGAGAAACA 1304  
DB 400 GluGlyAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer 417  
QY 1305 TGCTGTGATATTGCCACATGTAGATTTTAAACCCGTTCAAACTGTGCTGAAGACCATGC 1364  
DB 418 CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyrLysGlyLeuCys 437  
QY 1365 TCGAAAACTGCTCTATTATTCATCAAAAGAAAGATGTGTAGGCTCT---TCCTTTGAAGAA 1421  
DB 438 CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu 457  
QY 1422 TGGACCTCCCTGATATTTGCAATGAGATCATCTGCATCATGCCAGAAACCACTATGTT 1481  
DB 458 CysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyProAspIleThrLeu 477

QY 1482 CAGACTGGGCATCCGTGTGGAGTCAATCAATGGAATCTGTATAGATGGAGTTTGTATGAGT 1541  
DB 478 IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGlyAspCysHisAsp 497  
QY 1542 GGGGATAAACAATGTACACAGACACACATGGCAAGAAAGTAGAGTTTGGCCCTTCAGATGT 1601  
DB 498 LeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAlaProPheAlaCys 517  
QY 1602 TATTCTACCTTAATTAATCAAGACTGATGATCTCGAAACTGTGTGATAAGT---GATTCA 1658  
DB 518 TyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn 537  
QY 1659 GGATACACACAGTGTGAAGCTGACAAATCTGCAGTGGCGAAATAATTAATATGTAATGTGA 1718  
DB 538 LysTyrValPheCysGlyTrpArgAsnLeuIleCysGlyArgLeuValCysThrTyrPro 557  
QY 1719 GGTAAATTTTATTACAAATTTCCAAAGAGCCACTATTATTATGCAACATTAAGTGGACAT 1778  
DB 558 ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrAlaPheValArgAspSer 577  
QY 1779 CTCTGCATGCTGTGGAAATTTGCCAGTGATCATGCACAGCCCAAAAGATGTGGATAAAA 1838  
DB 578 ValCysIleThrValAspTyrLysLeuProArgThrValProAspProLeuAlaValLys 597  
QY 1839 GATGAACTTCTGCTGTTCAAAATAGGTTTCAGGAAATCAAGATGTGTGAGTTCTTCA 1898  
DB 598 AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg 617  
QY 1899 TACTTG-----GGTTATGATTGTACTACTACAAAATGCAATGATAGAGTGTATGC 1949  
DB 618 IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys 636  
QY 1950 AATAACAAAAGCACTGTCTACTGTAGTCTTCATATTACTCCAGATGTCTCAGTTCAA 2009  
DB 637 AspSerArgAsnLysCysHisCysSerProGlyTyrLysProAsnCysGlnIleArg 656  
QY 2010 TCAGATCTATGGCTGTGGGAGTATT-----GACAGTGCCAATTTTCCACCT 2057  
DB 657 Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle----- 671  
QY 2058 GTAGCTATACAGCCAGACTCCCTGAAAGCGCTACATTTGAGAACATTTTACCATTCCAAA 2117  
DB 672 -----MetGluArgAlaSerGlyLysThr 679  
QY 2118 CCAATGAGATGG-----CCATTTTCTTATTCATTCCTTTCTTTATTATTCTGTGTA 2171  
DB 680 GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal----- 696  
QY 2172 CTGATTGCTATATGTTGAAAGTTAATTTCCAAAGGAAAAAATGGAGAACTGAGGACTAT 2231  
DB 697 ---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu--- 714  
QY 2232 TCAAGCGATGACCAACTGAAAGTGAAGTGAACCTTAAGGG 2273  
DB 715 -----GluGluPheProSerSerGluSerLysSerGluGly 726  
RESULT 10  
ABU80783  
ID ABU80783 standard; protein; 787 AA.  
XX ABU80783;  
XX 23-JUN-2003 (first entry)  
XX Human PRO polypeptide #45.  
DE Human; PRO polypeptide; secreted and transmembrane protein;  
XX anti-PRO antibody; diagnostic assay; gene expression; tumour; cytostatic.  
KW Homo sapiens.  
OS US2003036635-A1.  
PN XX



QY	1542	GGGGATAAACAAATGTACACACACACATTTGGCGAAAGAGTAGAGTTGGCCCTTCAGAATGT	1601
Db	498	LeuAspAlaArgCysGluSerValPheGlyysGlySerArgAsnAlaProPheAlaCys	517
QY	1602	TATTCCTCACTTAAATCTCAAGACTGATGTATCTGGAAACTGTGTATAAGT--GATTCA	1658
Db	518	TyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn	537
QY	1659	GGATACACACAGTGGAGCTGACACATCTGCAGTGCAGCGCGAAATAATATGTAATAATGTGA	1718
Db	538	LysTyrValPheCysGlyTyrPArgAsnLeuIleCysGlyArgLeuValCysThrTyrPro	557
QY	1719	GGTAAATTTTATTACAAATCCAAAGAGCCACTATTATTTATGCCAATCAATAGTGGACAT	1778
Db	558	ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrAlaPheValArgAspSer	577
QY	1779	CTCTGCATTTGCTGGGAATTTGCCAGTGATCATCGACACAGCCAAAAGATGTGGATAAAA	1838
Db	578	ValCysIleThrValAspTyrLysLeuProArgThrValProAspProLeuAlaValLys	597
QY	1839	GATGGAACCTCTGTGGTTCAATAAGTTTGCAGGAATCAAGATGTGTGAGTCTTCTCA	1898
Db	598	AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg	617
QY	1899	TACTTG-----GGTTATGATGTACTACTGACAAATGCCAATGATAGAGGTGTATGC	1949
Db	618	IleIleLysAlaSerAlaHisValCys--SerGlnGlnCysSerGlyHisGlyValCys	636
QY	1950	AATAACAAAAGCACTGTCAGTCTAGTCTTCATATTTACTCGAGATTGCTCAGTCAAA	2009
Db	637	AspSerArgAsnLysCysHisCysSerProGlyTyrLysProAsnCysGlnIleArg	656
QY	2010	TCAGATCTATGGCCTGTGGGAGTATT-----CACAGTGCCAATTTTCCACCT	2057
Db	657	Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle-----	671
QY	2058	GTAGCTATACAGCCAGCACTCCCTGAAAGCGCTACATTGAGAACATTTTACCATTCCAAA	2117
Db	672	-----MetGluArgAlaSerGlyLysThr	679
QY	2118	CCAATGAGATGG-----CCATTTTCTTATTTCATTCCTTTCTTTATTATTTTCTGTGA	2171
Db	680	GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal	696
QY	2172	CTGATTGCTATAATGGTGAAGTTTAATTTCCAAAGGAAAAAATGGAGAACTGAGGACTAT	2231
Db	697	---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu---	714
QY	2232	TCAAGCGATGACCACTGAAAGTGAGATGAACCTAAAGGG	2273
Db	715	-----GluGluPheProSerSerGluSerLysSerGluGly	726

RESULT 11

RESULTS II  
ABO33749

ABO33743  
ID ABO33749 standard; protein: 787 AA.

XX  
XX

AC ABO33749;

XX

DT 17-SEP-2003 (first entry)

XX

DE Novel human secreted and t

**XX**

KW Human; secreted and transmembrane protein; PRO: cytosstatic;

KW antiarthritic; osteopathic; gene therapy: TNF-Agonist- $\alpha$ 1nba

KW chondrocyte stimulator; pericyte stimulator: fibroblast modulator; osteocyte stimulator; gene therapy; inf-agonist-ALPna

pharmaceutical; diagnostic; biosensor; bioreactor; tumour:

KW colon tumour; breast tumour; prostate tumour; rectal tumour

liver tumour; bone disorder: cartilage disorder: sports injury: prostate cancer; rectal cancer

KW  
arthritis; wound.  
arthritis; wound; sports injury; carriage disorder; sports injury; carriage disorder.

XX  
XX

OS Homo sapiens.

XX  
XX  
XXXXXXXXXXXX

PN	US2003045687-A1.
XX	
XX	
PD	06-MAR-2003.
XX	
XX	
PF	12-AUG-2002; 2002US-00218631.
XX	
XX	
PR	01-JUN-2001; 2001WO-US017800.
PR	29-JUN-2001; 2001WO-US021066.
PR	09-APR-2002; 2002US-00119480.
XX	
XX	
PA	(GETH ) GENENTECH INC.
XX	
PI	Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI	Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT;
XX	
XX	WPI; 2003-512315/48.
DR	N-PSDB; ACD68637.
DR	

PT New genes, and its encoded secreted and transmembrane polypeptides,  
PT useful for stimulating Tumor Necrosis Factor alpha, or chondrocyte or  
PT pericyte proliferation, especially for treating lung tumors, arthritis or  
PT wounds in a mammal

PS Claim 11; Fic 90; 314pp; English.

The invention describes an isolated nucleic acid molecule comprising a sequence with at least 80% identity to: (a) a nucleotide encoding any of 122 PRO (secreted and transmembrane) polypeptides whose sequences are fully defined in the specification; or (b) any of 122 nucleotide sequences having e.g. 4834, 2504 or 1759 bp fully defined in the specification; or the full length coding sequence of any these 122 nucleotide sequences. The PRO polypeptides or polynucleotides are useful as pharmaceuticals, diagnostics, biosensors or bioreactors. These are particularly useful for detecting tumours (e.g. lung tumour, colon tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour) in a mammal, for stimulating the release of TNF-alpha from human blood, for stimulating the proliferation or differentiation of chondrocyte cells, for stimulating proliferation of pericyte cells, or for modulating normal human dermal fibroblast proliferation. The PRO nucleic acid or polypeptide is also useful for treating tumours or various bone and/or cartilage disorders (e.g. sports injuries or arthritis), or wounds. The PRO polypeptides are useful in drug screening, particularly as targets for therapeutic intervention in these diseases, and in the diagnostic determination of the presence of these diseases. The PRO polypeptides are also useful as molecular weight markers, or for chromosome identification. The PRO genes are useful as hybridisation probes, or for screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene therapy, particularly for replacing a defective gene. This is the amino acid sequence of a novel human secreted and transmembrane PRO polypeptide

Sequence 787 AA:

Alignment Scores:

Pred. No.:	3,08e-127	Length:	787
Score:	1524.50	Matches:	310
Percent Similarity:	58.89%	Conservative:	134
Best Local Similarity:	41.11%	Mismatches:	259
Query Match:	31.79%	Indels:	51
DB:	6	Gaps:	18

US-10-054-683-18 (1-2640) x ABO33749 (1-787)

72 ATGTGGGCTCTGTTTCTCTCAGCGGGCTCGCGGGCTCGGATGCAC---AGTAATTTT 128  
" :  
4 LeutTpLeu-----LeuLeuAlaglyLeuCysglyLeuLeuAlaSerArgProGlyPhe 21  
:  
129 GATAGTTTACCTGTGCAAAATTACAGTTCGCGAGAAATACGGTCATAATAAAGGNAGA 188  
:  
22 GinAnSerLeuLeuGlnIleValileproGluYlsileGlnThrAsnThrAsnAspSer 41  
:  
189 ATTGAA-----TCGACGCCATCTACAAAATTGTAAATTGAAGGGGAAACCATATACT 239

Db 42 SerGluIleGluTyrGluGlnIleSerTyrIleIleProIleAspGluLysLeuTyrThr 61  
QY 240 GTGAATTAATGCAAAAAAATTTTACCCCATAAATTTTAGAGTTTACAGTTATAGTGGC 299  
Db 62 ValHisLeuLysGlnArgTyrPheLeuAlaAspAsnPheMetIleTyrLeuTyrAsn--- 80  
QY 300 ACAGGAATTAAGAACCATGACCTGACCAAGATTTTCAGAAATTTTCGCACATACCAAGGGTAT 359  
Db 81 GlnGlySerMetAsnThrTyrSerSerAspIleGlnThrGlnCysTyrTyrGlnGlyAsn 100  
QY 360 ATTGAAGGTTATCCAAATCTGTGGTGTATGTTAGCACATGTACTGGACTCAGGGCGTA 419  
Db 101 IleGluGlyTyrProAspSerMetValThrLeuSerThrCysSerGlyLeuArgGlyIle 120  
QY 420 CTACAGTTTCAAAATGTTAGTATGAATAAGAACGAGATGTTTCCCTTATATATAGAGAGAT 479  
Db 121 LeuGlnPheGluAsnValSerTyrGlyIleGluProLeuGluSerAlaValGluPheGln 140  
QY 480 CATGTTAATTTACCAAGTAACATAGAACGAGATGTTTCCCTTATATATAGAGAGAT 539  
Db 141 HisValLeuTyrLysLeuLysAsnGluAspAsnAspIleAlaIlePheIleAspArgSer 160  
QY 540 ATTGAATCAAGA-----GATCTCTCTTAAATTTACAAAGCCGACGACACGCAA 590  
Db 161 LeuLysGluGlnProMetAspAspAsnIlePheIleSerGluLysSerGluProAlaVal 180  
QY 591 GAT-----TTTCAAAAGTATAGAAATCATGTTATAGTTTCAAAACAAATTTGATAAT 644  
Db 181 ProAspLeuPheProLeuTyrLeuGluMetHisIleValValAspLysThrLeuTyrAsp 200  
QY 645 CATATGGGTCGATCAACCTGTGTGCTCAAAAAGTTTTCAGTTGATGATGAGTACG 704  
Db 201 TyrTrpGlySerAspSerMetIleValThrAsnLysValIleGluValGlyLeuAla 220  
QY 705 AATGCTATTCTTTTCATTATATATACAAATTTCTGCTTCATTGGAGCTTTGGATA 764  
Db 221 AsnSerMetPheThrGlnPheLysValThrIleValLeuSerSerLeuLeuLeuTrpSer 240  
QY 765 GATGAAATAAATAATGCAACCACTGAGAGAACTAATGAGTTATACACACATTTTAAAGA 824  
Db 241 AspGluAsnLysIleSerThrValGlyGluAlaAspGluLeuLysPheLeuGlu 260  
QY 825 TGAATAACATCTTATCTGTTTACGTCCTCATGATGGCATTTTACTTTGTTACAGA 884  
Db 261 TrpLysGlnSerTyrLeuAsnLeuAsgProHisAspIleAlaTyrLeuLeuIleTyrMet 280  
QY 885 GAAAGTCAAATTTATGTTGGTGAACCTTTCAAGGAGAGATGTTGATGCAAACTATGCA 944  
Db 281 AspTyrProArgTyrLeuGlyAlaValPheProGlyThrMetCysIleThrArgTyrSer 300  
QY 945 GGAGGTGTTGTTCTGCACCCCAAGAACCAATGCTGGAATCACTTGCAGTTATTTAGCT 1004  
Db 301 AlaGlyValAlaLeuTyrProLysGluIleThrLeuGluAlaPheAlaValIleValThr 320  
QY 1005 CAATTAATGACCTTAGTATGGGATCATCTTATGATGACATTAACAAATGCCAGTCTCA 1064  
Db 321 GlnMetLeuAlaLeuSerLeuGlyIleSerTyrAspAspProLysLysCysGlnCysSer 340  
QY 1065 GGAGCTGTCGATTAATGATCAGAACCAATTCATTTTCAGTGGTGTGAGATCTTTAGT 1124  
Db 341 GluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValLysThrPheSer 360  
QY 1125 AACTGCAGCTTCAAGACTTTGCACATTTTATTTTCAAGCAGAAAGTCCCAAGTCTTTCAC 1184  
Db 361 SerCysSerLeuArgSerPheGlnAsnPheIleSerAsnValGlyValLysCysLeuGln 380  
QY 1185 AATCAGCTCGCTTAGATCTCTTTTCAAAACAGCAAGCATGTTGTGTAATGCAAAAGCTG 1244  
Db 381 AsnLysProGlnMetGlnLys---LysSerProLysProValCysGlyAsnGlyArgLeu 399  
QY 1245 GAAGCAGGAGAGAGATGTTGATGCTGGGACTGAACAGGATTTGCTTATTTGGAGAAACA 1304  
Db 400 GluGlyAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer 417

RESULT 12

ABU82092

ID ABU82092 standard; protein; 787 AA.

XX

QY 1305 TGCTGTGATATTGCCACATGTAGATTAAAGCCGGTTCAAACTGTGCTGTAAGGACCAATGC 1364  
Db 418 CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyrLysGlyLeuCys 437  
QY 1365 TGGCAAAACTGCTATTATATGTCAAAGAAAGAAAGTATGTAGGCT---TCCTTTGAAGAA 1421  
Db 438 CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu 457  
QY 1422 TGGGACTCCCTGAATATTGCAATGATCATCTGCATCATCTGCCAGAAACCACTATGTT 1481  
Db 458 CysAspIleAlaGluAsnCysGlnGlySerSerProGluCysGlyProAspIleThrLeu 477  
QY 1482 CAGACTGGGCATCCGTGTGGACTGAATCAATCGATCTGTATATAGATGGAGTTTGTAGT 1541  
Db 478 IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGlyAspCysHisAsp 497  
QY 1542 GGGGATAAACAATGTACACACACATTTGGCAAGAAGTAGAGTTTGGCCCTTCAGATGT 1601  
Db 498 LeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAlaProPheAlaCys 517  
QY 1602 TATTCTCACCTTAATTTCAAGACTGATGTATCTGGAACCTGTGCTATAAGT---GATTCA 1658  
Db 518 TyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn 537  
QY 1659 GGATACACACAGTGTGAAGCTGCAATCTGCAGTGGGAAAATTAATATGTAAATATGTA 1718  
Db 538 LysTyrValPheCysGlyTrpArgAsnLeuIleCysGlyArgLeuValCysThrTyrPro 557  
QY 1719 GGTAATTTTATTACAAATTTCCAGAGCCACTATTATTATTCGCAACATAAGTGGNAT 1778  
Db 558 ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrAlaPheValArgAspSer 577  
QY 1779 CTCTGATTTGCTGCAATTTTGCAGTGTATCATGCAGACAGCCAAAGAGTGTGATAAAA 1838  
Db 578 ValCysIleThrValAspTyrLysLeuProArgThrValProAspProLeuAlaValLys 597  
QY 1839 GATGGAACCTTTTGTGTTTCAATAAGAGTTTGCAGGAATCAAGATGTGTGAGTTCTTCA 1898  
Db 598 AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg 617  
QY 1899 TACTTG-----GTTATGATGTACTACTGCAAAATGCAATGCAATGATAGAGTGTATGC 1949  
Db 618 IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys 636  
QY 1950 AATAACAAAAGCAGCTGCTAGTCTGCTTTCATATTTTACCTCCAGATTCCTCAGTTCAA 2009  
Db 637 AspSerArgAsnLysCysHisCysSerProGlyTyrLysProAsnCysGlnIleArg 656  
QY 2010 TCAGATCTATGGCTGTGGAGTATT-----GACAGTGGCAATTTTCCACCT 2057  
Db 657 Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle----- 671  
QY 2058 GTAGCTATACCAGCCAGACTCCCTGAAAGCGCTACATTGAGAACACATTACCATTCCAAA 2117  
Db 672 -----MetGluArgAlaSerGlyLysThr 679  
QY 2118 CCAATCAGATGG-----CCATTTTCTTATTCTTCTTCTTCTTCTTCTTCTGTA 2171  
Db 680 GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal----- 696  
QY 2172 CTGATTGCTATATGTTGAAGTTAATTTCCAAAGGAAAAAATGAGAACTGAGGACTAT 2231  
Db 697 ---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu--- 714  
QY 2232 TCAAGCGATGAGCAACTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2273  
Db 715 -----GluGluPheProSerSerGluSerLysSerGluGly 726



QY	1245	GAACGACGAGGAGTGTGACTGTGGGACTGAACACGAGATTGTGCCCTTATTGTGGAGAAACA	1304
DB	400	GlulysAsnGluileCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer	417
QY	1305	TGCTGTGATATGGCCACATGTAGATTAAAGCCGGTTCAAACCTGTCTGCTGAAGGACCATTGC	1364
DB	418	CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyLysGlyLeuLysCys	437
QY	1365	TGCGAAACATGCTCTATTATTGTCTCAAAAGAAGATGTGTAGGCT---TCCTTTGAGAA	1421
DB	438	CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu	457
QY	1422	TGGACCTCCTCGAATATTGCAATGGATCATCTGCATCATGCCCAGAAAAACCATGATTGT	1481
DB	458	CysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyProAspIleThrLeu	477
QY	1482	CAGACTGGGCATCGCTGTGGACTGAAATCAATGGATCTGTATAGATGAGTTTGTATGAGT	1541
DB	478	IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGlyAspCysHisAsp	497
QY	1542	GGGGATAAACAAATGTACAGACACATTTGGCAAGAAAGTAGAGTTTGGCCCTTCAGAAATGT	1601
DB	498	LeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAlaProPheAlaCys	517
QY	1602	TATTCTCACCTTAATTCAAAGACTGATGTATCTGGAACCTGTGGTATAAGT---GATTCA	1658
DB	518	TyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn	537
QY	1659	GGATACACACAGTGTGAAGCTGACAACTCTGCAGTGGCGAAATTAATATATGTAAATATGTA	1718
DB	538	LysTyrValPheCysGlyTyrPArgAsnLeuIleCysGlyArgLeuValCysThrTyrPro	557
QY	1719	GGTAAATTTTATPACAAATTCCAAGAGCCACTATTATTATGGCAACATAAGTGGACAT	1778
DB	558	ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrAlaPheValArgAspSer	577
QY	1779	CTCTGCATCTCTGGGAATTTGCCAGTGATCATGCACACGCCCAAGATGTGGATAAAA	1838
DB	578	ValCysIleThrValAspTyrLysLeuProArgThrValProAspProLeuAlaValLys	597
QY	1839	GATGGAACTTCTTGTGGTTCAATAAGCTTTGCAGGAATCAAGAGATGTGTGAGTTCTTCA	1898
DB	598	AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg	617
QY	1899	TACTTG-----GGTTAGATTGTACTGACAAATGCAATGATAGAGGTGATGC	1949
DB	618	IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys	636
QY	1950	AATAACAAAAGCACTCTCACTAGTGCCTTCATATTACTCCAGATTCTCGATTCAA	2009
DB	637	AspSerArgAsnLysCysHisCysSerProGlyTyrLysProProAsnCysGlnIleArg	656
QY	2010	TCAGATCTATGGCTGTGGGAGTATT-----GACATGGCAATTTTCCACCT	2057
DB	657	Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle-----	671
QY	2058	GTAGCTATACGAGCCAGCTCCCTGGAAGGGCTACATTGAGAACATTTACCATTCCAA	2117
DB	672	-----MetGluArgAlaSerGlyLysThr	679
QY	2118	CCAATGAGATGG-----CCATTTTCTTATTCAATCTCTTCTTATTATTCTTGTA	2171
DB	680	GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal-----	696
QY	2172	CTGATTCCTATAATGGTGAAGTTAAATTTCCAAGGAAAAAATGGAGAACTGAGGACTAT	2231
DB	697	---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysIleTrpPheAlaLysGlu---	714
QY	2232	TCAAGCGATGAGCAACTGAAAGTGAAGTGAAGTGAAGCTTAAGGG	2273
DB	715	-----GluGluPheProSerSerGluSerLysSerGluGly	726

RESULT 13	
ABU72272	
ID	ABU72272 standard; protein; 787 AA.
XX	
AC	ABU72272;
XX	
DT	06-NOV-2003 (first entry)
XX	
DE	Human PRO21340 protein.
XX	
KW	PRO; proliferation; pericyte cell; TNF-alpha; blood; chondrocyte;
KW	differentiation; dermal fibroblast; tumour; gene therapy; cytostatic.
XX	
OS	Homo sapiens.
XX	
FN	US2003050448-A1.
XX	
PD	13-MAR-2003.
XX	
FF	28-AUG-2002; 2002US-00230414.
XX	
PR	01-JUN-2001; 2001WO-US017800.
PR	29-JUN-2001; 2001WO-US021066.
PR	09-APR-2002; 2002US-00119480.
XX	
FA	(GETH ) GENENTECH INC.
XX	
PI	Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI	Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX	
DR	WPI: 2003-521818/49.
DR	N-PSDB; ABT44270.
XX	
FT	New nucleic acid encoding for a PRO protein, useful for the manufacture
FT	of a medicament for diagnosing or treating tumors or for measuring or
FT	detecting expression of an associated gene.
XX	
PS	Claim 11; Fig 90; 315pp; English.
XX	
CC	The invention relates to a novel isolated nucleic acid encoding a fully
CC	defined PRO polypeptide. The molecules of the invention may be useful for
CC	stimulating proliferation or gene expression in pericyte cells or the
CC	release of TNF-alpha from human blood. Other possible uses include the
CC	stimulation or inhibition of chondrocyte proliferation or
CC	differentiation, the stimulation of human dermal fibroblast cell
CC	proliferation and the detection of the presence of a tumour within a
CC	mammal. Furthermore, the nucleic acid may be useful for the manufacture
CC	of a medicament for diagnosing or treating a tumour within a mammal or
CC	for measuring or detecting the expression of an associated gene, as well
CC	as during gene therapy. The current sequence is that of the human PRO
CC	protein of the invention
XX	
SQ	Sequence 787 AA;
	Alignment Scores:
Pred. No.:	3,08e-127 Length: 787
Score:	1524.50 Matches: 310
Percent Similarity:	58.89% Conservative: 134
Best Local Similarity:	41.11% Mismatches: 259
Query Match:	31.79% Indels: 51
DB:	6 Gaps: 18
US-10-054-683-18 (1-2640) x ABU72272 (1-787)	
Qy	72 ATGTGGGTCTGTGTTCTGCTCAGCGGCTCGCGGGTGGCGATGGAC---AGTAATTTT 128
.	::: :::
Db	4 LeuTrpLeu-----LeuLeuAlaGlyLeuLeuAlaSerArgProGlyphe 21
Qy	129 GATAGTTTACCTGTGCAAAATACAGTTCCGAGAAATACGGTCAATAATAAGCAAGGA 188
.	::: :::
Db	22 GlnAsnSerLeuLeuGlnIleValIleProGlnIlySileGlnThrAsnThrAsnAspSer 41









```
Db 418 CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyrLysGlyLeuCys 437
QY 1365 TGGAAAGCTGCTCTATTTATGTCAGAAAGAAAGTGTAGGCCT---TCCTTTGAAGAA 1421
Db 438 CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu 457
QY 1422 TCGACCTCCCTGATATTTGCAATGATCATCTGTCATCATGCCAGAAACCACTATGTT 1481
Db 458 CysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyProAspIleThrLeu 477
QY 1482 CAGACTGGGCATCCGTGGACTGATCAATGATGATGATGATGATGATGATGATGATGAT 1541
Db 478 IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGlyAspCysHisAsp 497
QY 1542 GGGGATAAACAATGTACACACATTTGGCAAGAAGTAGATTTGGCCCTTCAGAAATG 1601
Db 498 LeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAlaProPheAlaCys 517
QY 1602 TATCTCACCTTAATTCAAAGACTGATGTATCTGGAAGCTGTGGTATAAGT---GATTCA 1658
Db 518 TyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspAsnAsn 537
QY 1659 GGATACACACAGTGTGAAGCTGACAAATCTGCAGTCGCGGAAATTAATGATGATGATGAT 1718
Db 538 LysTyrValPheCysGlyTyrPargAsnLeuIleCysGlyArgLeuValCysThrTyrPro 557
QY 1719 GGTAATATTTTATTACAAATTCAGAGCCACTATTATTATTTATGCAACATTAAGTGCAT 1778
Db 558 ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrAlaPheValArgAspSer 577
QY 1779 CTCTGCATGCTGTGGAAATTTGCCAGTGATCATGCAGACAGCCAAAGATGTGGATAAAA 1838
Db 578 ValCysIleThrValAspTyrLysLeuProArgThrValProAspProLeuAlaValLys 597
QY 1839 GATGAACTCTTGGGTTTCAATTAAGGTTTGCAGGAATCAAGATGTGTGAGTCTTCA 1898
Db 598 AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg 617
QY 1899 TACTTG-----GGTTATGATTGTACTACTGACAATCAATGATAGAGGTGATGC 1949
Db 618 IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys 636
QY 1950 AATCAACAAAGCACTGCTACTGTAGTCTTCAATTTACTTCAGATTGCTCAGTTCAA 2009
Db 637 AspSerArgAsnLysCysHisCysSerProGlyTyrLysProAsnCysGlnIleArg 656
QY 2010 TCAGATCTATGCGCTGGTGGGAGTATT-----GACATGGCAATTTTCCACCT 2057
Db 657 Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle----- 671
QY 2058 GTAGCTATACAGCCAGACTCCCTGAAAGGCGCTACATTTGAGAACATTTTACCATTCCAAA 2117
Db 672 -----MetGluArgAlaSerGlyLysThr 679
QY 2118 CCAATGATAGTG-----CCATTTTCTTATCTATCTTCTTTTATTTATTTCTGTGTA 2171
Db 680 GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal----- 696
QY 2172 CTGATTGCTATAATGGTGAAGTAAATTTCCAAAGGAAAAAATGGAGAACTGAGGACTAT 2231
Db 697 ---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysIleTrpPheAlaLysGlu--- 714
QY 2232 TCAAGCATGAGCACTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2273
Db 715 -----GluGluPheProSerSerGluSerLysSerGluGly 726
RESULT 15
ID ABO34295
XX ABO34295 standard; protein; 787 AA.
AC ABO34295;
XX
XX DT 19-SEP-2003 (first entry)
```

```
XX Human secreted/transmembrane polypeptide PRO 21340.
DE Human; chondrocyte stimulation; TNF-alpha stimulation; gene therapy;
KW human dermal fibroblast stimulation; tumour; tissue typing;
KW affinity purification.
XX
OS Homo sapiens.
XX US2003044934-A1.
XX 06-MAR-2003.
XX 28-AUG-2002; 2002US-00230338.
XX 01-JUN-2001; 2001WO-US017800.
XX 29-JUN-2001; 2001WO-US021066.
XX 09-APR-2002; 2002US-00119480.
XX (GETH ) GENENTECH INC.
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI; 2003-492274/46.
XX N-PSDB; ACD82220.
XX New transmembrane polypeptides and nucleic acids encoding the
XX polypeptides, useful in gene therapy, in chromosome identification, as
XX chromosome markers, or in generating probes.
XX Claim 19; Fig 90; 315pp; English.
XX The invention relates to an isolated nucleic acid encoding a PRO
XX polypeptide. Nucleic acids that encode PRO can be used to generate either
XX transgenic animals or knock-out animals useful in developing and
XX screening of therapeutically useful reagents. The nucleic acids may also
XX be used in gene therapy for replacing defective gene, in chromosome
XX identification, as chromosome markers, or in generating probes to isolate
XX full length PRO cDNA. The PRO polypeptides are useful for chondrocyte
XX stimulation, TNF-alpha stimulation, human dermal fibroblasts stimulation
XX and for detecting the presence of tumour in an animal. The PRO
XX polypeptides are useful as molecular markers for protein electrophoresis
XX and the isolated nucleic acids may be used for recombinantly expressing
XX those markers. The PRO polypeptides and nucleic acids may also be used in
XX tissue typing. Anti-PRO antibodies are useful in diagnostic assays for
XX PRO and in affinity purification of PRO from recombinant cell culture or
XX natural sources. The present sequence represents the amino acid sequence
XX of a human secreted/transmembrane PRO polypeptide
XX SQ Sequence 787 AA;
Alignment Scores:
Pred. No.: 3,08e-127 Length: 787
Score: 1524.50 Matches: 310
Percent Similarity: 58.89% Conservative: 134
Best Local Similarity: 41.11% Mismatches: 259
Query Match: 31.79% Indels: 51
DB: 6 Gaps: 18
US-10-054-683-18 (1-2640) x ABO34295 (1-787)
QY 72 ATGTGGTCTTCTTCTCTGCTCAGCGGCTCGCGGGCTCGGATGAC---AGTAATTTT 128
Db 4 LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe 21.
QY 129 GATAGTTTACCTGTCACAAATTACAGTTCCGGAGAAATACGCTCAATAATAAGGAAGA 188
Db 22 GlnAsnSerLeuLeuGlnIleValIleProGluLysIleGlnThrAsnThrAsnAspSer 41
QY 189 ATTGAA-----TCCGAGGCATCTCCACAAATTTGTAATTAAGAGGAACCATATACT 239
Db 42 SerGluIleGluTyrGluGluIleSerTyrIleIleProfileAepGluLysLeuTyrThr 61
```



*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 10, 2005, 21:32:40 ; Search time 90.5 Seconds  
(without alignments)  
5613.530 Million cell updates/sec

Title: US-10-054-683-18

Perfect score: 4796

Sequence: 1 catctcgacttccaactgc.....actgttacatttcaaaaaa 2640

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-O=/cpn2\_1/USPTO\_spool\_p/US10054683/runat\_10012005\_174510\_15504/app\_query.fasta\_1.2823  
-DB=PIR79 -QPMF=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10054683 @CCN 1 1 135 @runat\_10012005\_174510\_15504 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	3984	83.1	734	JC4861	fertilin beta cha
2	3687.5	76.9	735	G02937	fertilin beta - cr
3	1384.5	28.9	756	S47656	tMDC II protein -
4	1365.5	28.5	736	S47645	tMDC I protein - c
5	1292	26.9	823	S18968	cyritestin precurs
6	1288	26.9	777	I48100	ADAM 5 protein pre
7	1186	24.3	357	S23403	sperm surface prot
8	1101	23.0	655	JC7850	disintegrin and me
9	1033.5	21.5	825	S55060	fertilin alpha-II
10	1022.5	21.3	905	S55059	fertilin alpha-I -
11	1016.5	21.2	660	S71549	metalloproteinase
12	857	17.9	903	S60257	meltrin alpha - mo
13	795	16.6	600	I49281	fertilin alpha pre
14	766	16.0	826	A60385	monocyte surface a

15	762	15.9	609	2	S55270	catrocollastatin p
16	732	15.3	732	2	I52361	testicular metallo
17	729	15.2	617	2	S48160	metalloproteinase
18	723	15.1	814	2	G02390	disintegrin-like m
19	719.5	15.0	571	2	S24789	jararhagin C precu
20	712.5	14.9	735	2	I48101	ADAM 6 protein pre
21	710.5	14.8	616	2	A55796	ecarin precursor
22	699	14.6	776	2	S28258	androgen-regulated
23	696.5	14.5	789	2	S28259	androgen-regulated
24	695	14.5	670	2	I65967	disintegrin-like m
25	692.5	14.4	713	2	I65253	disintegrin-like t
26	663	13.8	610	2	JC7530	vascular apoptosis
27	653.5	13.6	610	2	JC8056	halyase - Glycidu
28	641	13.4	952	2	T18900	disintegrin and me
29	616.5	12.9	429	2	A42972	coagulation factor
30	614.5	12.8	416	2	A37877	hemorrhagic protei
31	606	12.6	419	2	S41607	atrolysin A (EC 3.
32	594.5	12.4	419	2	A59414	metalloproteinase
33	584.5	12.2	549	2	S48169	metalloproteinase
34	568	11.8	473	2	I49283	ADAM 4 protein pre
35	561.5	11.7	524	2	S38539	disintegrin-like m
36	553	11.5	478	2	A43296	atrolysin E (EC 3.
37	535	11.2	484	2	JC8020	metalloproteinase-
38	533	11.1	480	1	A30065	trigramin precurs
39	530.5	11.1	481	2	JC4342	fibrinolytic prote
40	530.5	11.1	481	2	S43125	trimucin precursor
41	501	10.4	478	2	JQ1301	hemorrhagic protei
42	486.5	10.1	1042	2	T26644	hypothetical prote
43	467	9.7	478	2	JC4880	fibrinolytic metal
44	393.5	8.2	414	2	S41608	atrolysin B (EC 3.
45	392.5	8.2	411	1	HY5NFA	fibrolase (EC 3.4.

ALIGNMENTS

RESULT 1

JC4861  
fertilin beta chain - human  
C:Species: Homo sapiens (man)  
C>Date: 15-Aug-1996 #sequence\_revision 18-Oct-1996 #text\_change 09-Jul-2004  
C:Accession: JC4861  
R:Gupta, S.K.; Alves, K.; O'Neil Palladino, L.; Mark, G.E.; Hollis, G.F.  
Biochem. Biophys. Res. Commun. 224, 318-326, 1996  
A:Title: Molecular cloning of the human fertilin beta subunit.  
A:Reference number: JC4861; MUID:96295488; PMID:8702389  
A:Accession: JC4861  
A:Molecule type: mRNA  
A:Residues: 1-734 <CUP>  
A:Cross-references: UNIPROT:Q99965; GB:U38805; NID:G4151118; PIDN:AD04206.1; PID:G415111  
C:Comment: This protein is an integral sperm membrane glycoprotein, and plays a role in  
C:Superfamily: mouse meltrin alpha; disintegrin homology  
C:Keywords: glycoprotein; integrin binding; transmembrane protein  
F:382-734/Product: fertilin beta chain #status predicted <NAT>  
F:382-467/Domain: disintegrin homology <DIS>  
F:448-450/Region: integrin binding #status predicted  
F:686-708/Domain: transmembrane #status predicted <TMW>  
F:121,219,352,458,565/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:			
Pred. No.:	1,32e-286	Length:	734
Score:	3984.00	Matches:	734
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	83.07%	Indels:	0
DB:	2	Gaps:	0

US-10-054-683-18 (1-2640) x JC4861 (1-734)

QY	72	ATGTGGTCTTGTTCCTGCTCAGCGGCTCGCGGCTCGGATGGACAGTATTTTGTAT	131
Db	1	MetTrpValLeuPheLeuLeuSerGlyLeuGlyLeuArgMetAspSerAanPheAap	20
QY	132	AGTTTACCTGTGCMAATTACGTTCCGGAGAAATACGGTCAATATAAAGGAAGGAAATT	191

Db	21	SerLeuProValGlnIleThrValProGluLysIleArgSerIleIleLysGluGlyIle	40	Db	381	LysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCysGly	400
Qy	192	GAATCCGAGCATCTACAAAATTGTAATTGAAGGAGAAACCATATATCTGTGAATTTAATG	251	Qy	1272	ACTGAAAGAGATTGGCCCTTATTGGGAAACATGCTGTGATATTGCCACATGTAGATT	1331
Db	41	GluSerGlnAlaSerTyrLysIleValIleGluGlyLysProTyrThrValAsnLeuMet	60	Db	401	ThrGluGlnAspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPhe	420
Qy	252	CAAAAAACCTTTTACCCCATAAATTTAGAGTTTACAGTTATAGTGGCAGGAAATTATG	311	Qy	1332	AAAGCCGGTTCAAACTGCTGTAAGGACCATGCTGCGAAAACTGTCTATTATTTATGTC	1391
Db	61	GlnLysAsnPheLeuProHisAsnPheArgValTyrSerTyrSerGlyThrGlyIleMet	80	Db	421	LysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLys	440
Qy	312	AAACCACTTCAGCAAGATTTTCAGATTTCTGCCACTACCAAGGCTATATTGAAGTTAT	371	Qy	1392	GAAAGAAATGTGTAGCCCTTCCTTGAAGAATGCGACCTCCCTCGAATATATGCAATGATCA	1451
Db	81	LysProLeuAspGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGlyTyr	100	Db	441	GluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGlySer	460
Qy	372	CCAAATCTGTGTGTGTTAGTACATGTACTGAGCTCAGGGCGCTACTACAGTTTGAA	431	Qy	1452	TCGTGATCTGATAGATGGAGTTTGTATGATGGGATAAACAATGTACAGACACATTTGGC	1511
Db	101	ProLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGlu	120	Db	461	SerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGln	480
Qy	432	AATGTTAGTTATGGAATAGAACCCCTGGAGTCTTTCAGTTGGCTTGAACATGTAATTAC	491	Qy	1512	TGGATCTGTATAGATGGAGTTTGTATGATGGGATAAACAATGTACAGACACATTTGGC	1571
Db	121	AsnValSerTyrGlyIleGluProLeuGluSerSerValGlyPheGluHisValIleTyr	140	Db	481	TrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGly	500
Qy	492	CAAGTAAACATAGAAGCAGATGTTTCTTATATATGAGAGGATATTGAATCAAGA	551	Qy	1572	AAAGAAATAGATTGGCCCTTCAGAAATGTTATTTCTCACCTTAATTCAGAACTCATGTA	1631
Db	141	GlnValLysHisLysLysAlaAspValSerLeuTyrAsnGlnLysAspIleGluSerArg	160	Db	501	LysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspVal	520
Qy	552	GATCTGTCTTTAAATACAAAGCGCAGACCCAGCAAGATTTTGCAGATATATAGAA	611	Qy	1632	TCGTGAAACTGTGGTATAGTATTGATTAAGTATGATTAAGTATTAAGTATTAAGTATTA	1691
Db	161	AspLeuSerPheLysLeuGlnSerAlaGluProGlnAspPheAlaLysTyrIleGlu	180	Db	521	SerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGln	540
Qy	612	ATGCATGTTATAGTTGAAACAAATGCTATATATCATATGCGGTCTGATACAACTGTGC	671	Qy	1692	TGCGGAAATTAATATGTAATATGTAATATGTAATATTAATTAATTAATTAATTAATTA	1751
Db	181	MetHisValIleValGluLysGlnLeuTyrAsnHisMetGlySerAspThrThrValVal	200	Db	541	CysGlyLysLeuIleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAlaThr	560
Qy	672	GCTCAAAAGTTTCCAGTTGATTGGATTGACGAATGCTATTTTGTTCATTTAATATT	731	Qy	1752	ATTATTTATGCCAACATAAGTGGACATCTCTGATGCTGTGCGAATTTGCCAGTGCATCAT	1811
Db	201	AlaGlnLysValPheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsnIle	220	Db	561	IleIleTyrAlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHis	580
Qy	732	ACAATTTATCTGCTTCATTTGGAGCTTTGGATAGATGAAATAAATATGCAACCTGGA	791	Qy	1812	GCAGACAGCAAAAGATGTGGATAAAGATGAAACTTCTTGTGGTTCAAAATGAAGTTTC	1871
Db	221	ThrIleIleLeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThrThrGly	240	Db	581	AlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCys	600
Qy	792	GAACCTAATGAGTTTATACACACATTTTAAAGATGGAACATCTTATCTTTTACGT	851	Qy	1872	AGGAATCAAGATGTGTGAGTTCTTCATCTTGGGTTATGATTTGATCTACTGACAAATGC	1931
Db	241	GluAlaAsnGluLeuLeuHisThrPheLeuArgTrpLysThrSerTyrLeuValLeuArg	260	Db	601	ArgAsnGlnArgCysValSerSerSerTyrLeuGlyTyrAspCysThrThrAspLysCys	620
Qy	852	CCTCATGATGTGGCATTTTACTTGTTTTACAGAAAGTCAAAATATGTTGGTGCAACC	911	Qy	1932	AATGATAGAGCTGTATGCAATTAACAAAGACACTGTCTGCTAGTGTCTTCAATTTACCT	1991
Db	261	ProHisAspValAlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAlaThr	280	Db	621	AsnAspArgGlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeuPro	640
Qy	912	TTTCAAGGAAGATGTGTGATGCAAACTATGCGAGAGTGTGTTCTGCACCCACGAACC	971	Qy	1992	CCAGATTGCTCAGTTCAATCAGATCTATGCGCTGTGGAGTATTGACAGTGGCAATTTT	2051
Db	281	PheGlnGlyLysMetCysAspAlaAsnTyrAlaGlyValValLeuHisProArgThr	300	Db	641	ProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsnPhe	660
Qy	972	ATAAGTCTGGAATCACTTGCAGTTATTTTACTCAATTTAGTCCTTAGTATGGGATC	1031	Qy	2052	CCACCTGTAGCTATACAGCCAGACTCCCTGAAAGCGCTACATTTGAGAACATTTACCAT	2111
Db	301	IleSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGlyIle	320	Db	661	ProProValAlaIleProAlaArgLeuProGluArgArgTyrIleGluAsnIleTyrHis	680
Qy	1032	ACTTATGATGACATTAACAATGCGAGTGTCTCAGAGCTGTCTGATTTGAAATCCAGAA	1091	Qy	2112	TCCAAACCAATGAGATGGCCATTTTCTTATTCTTCTTCTTCTTCTTCTTCTTCTGTA	2171
Db	321	ThrTyrAspAspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnProGlu	340	Db	681	SerLysProMetArgTrpProPhePheLeuPheIleProPhePheIleIlePheCysVal	700
Qy	1092	GCAATTCATTTTCAGTGGTGAAGATCTTTAGTAATCTGACCTTCGAGACTTTGGACAT	1151	Qy	2172	CTGATTGCTATAATGTTGAAGTTAATTTCCAAAGGAAAAAATGAGAACTGAGGACTAT	2231
Db	341	AlaIleHisPheSerGlyValLysIlePheSerAsnCysSerPheGluAspPheAlaHis	360	Db	701	LeuIleAlaIleMetValLysValAsnPheGlnArgLysLysTrpArgThrGluAspTyr	720
Qy	1152	TTTATTTCAAGCAGAGTCCAGGTCTTTCACATCAGCTCGCTTAGATCCTTTTTC	1211	Qy	2232	TCAGCGATCAGCAACCTGAACTGAGCTGAACTTAAAGG 2273	
Db	361	PheIleSerLysGlnLysSerGlnCysLeuHisLeuGlnProArgLeuAspProPhePhe	380	Db	721	SerSerAspGluGlnProGluSerGluSerGluProLysGly 734	
Qy	1212	AAACAGCAGCAGTGTGTGGTAAATCCAAAGCTGGAAGCAGGAGGATGTGACTGTGGG	1271				

C;Species: Macaca fascicularis (crab-eating macaque)  
 C;Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 09-Jul-2004  
 C;Accession: G02937; S55061  
 R;Ramrao, C.S.; Myles, D.G.; White, J.M.; Primakoff, P.  
 submitted to the EMBL Data Library, August 1995  
 A;Reference number: G12615  
 A;Accession: G02937  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-735 <RAM>  
 A;Cross-references: UNIPROT:Q28478; EMBL:U33959; NID:g998339; PID:g998340  
 R;Perry, A.C.P.; Gichuhi, P.M.; Jones, R.; Hall, L.  
 Biochem. J. 307, 843-850, 1995  
 A;Title: Cloning and analysis of monkey fertilin reveals novel alpha subunit isoforms.  
 A;Reference number: S55059; MUID:95260313; PMID:7741716  
 C;Superfamily: mouse meltrin alpha; disintegrin homology  
 F;383-468/Domain: disintegrin homology <DIS>

Alignment Scores:  
 Pred. No.: 1.le-264 Length: 735  
 Score: 3687.50 Matches: 671  
 Percent Similarity: 95.51% Conservative: 31  
 Best Local Similarity: 91.29% Mismatches: 32  
 Query Match: 76.89% Indels: 1  
 DB: 2 Gaps: 1

US-10-054-683-18 (1-2640) x G02937 (1-735)

Qy	72	ATGTGG---GTCCTGTTCTGCTCAGCGGGCTCGGGGCTCGGATGGACGATTAATTTT	128
Db	1	MetTrrpValIleuPheLeuSerGlyLeuGlyLeuThrPheAspSerAspPhe	20
Qy	129	GATGATTTACCTGTGCAAAATTACAGTTCCGGAGAAAATACGGTCAATAATAAAGGAGGA	188
Db	21	AspSerLeuProValGlnIleThrValProGluLysIleArgSerIleIleLysGluGlu	40
Qy	189	ATTGAATCGGAGCATCCTACAAATTTGTAATTGAAGGGAACCATATACCTGTAATTTA	248
Db	41	IleGluSerGlnValSerTyrLysIleValIleGluGlyLysProTyrThrAlaAsnLeu	60
Qy	249	ATGCAAAAACCTTTTACCCCATATTTAGAGTTTACAGTTTATAGTGGCACAGGAAT	308
Db	61	MetGlnLysAsnPheLeuSerHisAsnPheArgValTyrSerTyrAsnGlyThrGlyIle	80
Qy	309	ATGAAACCACTTGACCAAGATTTTCAGAAATTTCTGCCACTACCAAGGGTATATTGAAGT	368
Db	81	MetLysProLeuAspGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGly	100
Qy	369	TATCCAAATCTGTGTGATGTTAGACATGATGACTACAGGCTAGGGCGTACTACAGTTT	428
Db	101	TyrProLysSerValAlaMetValSerThrCysThrGlyLeuArgGlyLeuLeuGlnPhe	120
Qy	429	GAAATGTTAGTTATGGAATAGAACCCCTCGAGTCTTCAGTTCGCTTGAACATGTAAT	488
Db	121	GluAsnValSerTyrGlyIleGluProLeuGluSerValGlyPheGluHisValIle	140
Qy	489	TACCAAGTAAACATAGAAGCAGATGTTTCTCTATATATATGAGAGGATATTGAATCA	548
Db	141	TyrGlnValIlyshIlyLysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSer	160
Qy	549	AGAGATCTGCTCTTAAATTTACAAAGCGGACAGCCACAGCAAGATTTTGAAGTATATA	608
Db	161	ArgAspLeuSerPheLysLeuGlnSerIleGluProGlnLysAspPheAlaLysTyrIle	180
Qy	609	GAAATGATCTTATAGTTGAAAACCAATTTGTAATCATATGGTCTGTATACACTGTT	668
Db	181	GluMetHisValValGluLysGlnLeuTyrAsnHisMetGlySerGlyThrThrVal	200

Qy	669	GTGCTCAAAAAGCTTTTCCAGTTGATGATGAGTTCAGAACTGCTATTTTGTTCATTAAT	728
Db	201	ValThrGlnLysIlePheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerLeuAsn	220
Qy	729	ATTACAAATATTCTGCTTCCATTTGGAGCTTTGGATAGATGAAATATAAATTTGCAACACT	788
Db	221	IleThrValIleLeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThrThr	240
Qy	789	GGAGAAGCTAATGAGTTATTACACACATTTTAAAGATGGAACATCTTATCTTGTTTTA	848
Db	241	GlyAspAlaLysGluLeuLeuHisThrPheLeuArgTrrLysArgSerTyrLeuValLeu	260
Qy	849	CGTCTCTATGATGTGGCATTTTACTTTGTTCAGAGAAAGTCAATATGTTGTGTGCA	908
Db	261	ArgProHisAspValAlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAla	280
Qy	909	ACCTTTCAAGGGAAGATGTGTGATGCAAACTATGCAGGAGGTCTGTCTGCACCCGAGA	968
Db	281	ThrPheGlnGlyLysMetCysAspAlaAsnTyrAlaGlyValLeuLeuHisProArg	300
Qy	969	ACCATAAGTCTGGAATCAGTTCAGTATTATTAGCTCAATATTAGACCTTAGTATGGGG	1028
Db	301	ThrIleSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGly	320
Qy	1029	ATCATTATGATGACATTAACAATGCCAGTGTCTCAGGAGCTGTCTGCATATTGAATCCA	1088
Db	321	IleProTyrAspAspIleAsnGlnCysGlnCysSerAlaAlaValCysIleMetAsnPro	340
Qy	1089	GAAGCAATTCATTTAGTGTGTGAAGATCTTTAGTAACATGCAGCTTCGAAGACTTTGCA	1148
Db	341	GluAlaIleHisPheSerGlyValLysIlePheSerAsnCysSerIleGluAspPheAla	360
Qy	1149	CATTTTATTTCAAGCAGAAAGTCCCAAGTGTCTTCACAATCAGCTCGCTTAGATCTCTTT	1208
Db	361	HisPheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhe	380
Qy	1209	TTCAAA CAGCAAGCAGTGTGTGTAATGCAAGCTGGAAGCAGAGAGAGTGTGACTGT	1268
Db	381	PheLysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCys	400
Qy	1269	GGGACTGAACAGATTTGTCCTTATTGGAGAAACATGCTGTGATATTGCCACATGTAGA	1328
Db	401	GlyThrGlnGlnAsnCysPheLeuLeuGlyAlaLysCysCysAspThrAlaThrCysArg	420
Qy	1329	TTTAAAGCCGGTTTCAAACTGTCTGAAGGACCACTGCTGCCAAAACCTGCTTATTATGCA	1388
Db	421	PheLysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSer	440
Qy	1389	AAAGAAGAAATGTGTAGGCTTCTCTTTGAAGAATGCGACCTCCCTGAAATTTGCAATGGA	1448
Db	441	GlnGluArgValCysArgProSerPheAspGluCysAspLeuProGluTyrCysAsnGly	460
Qy	1449	TCATCTGCATCATGCCAGAAAACCACTATGTTTCAGACTGGGCTCGTGTGACTGAAT	1508
Db	461	ThrSerAlaSerCysProGluAsnHisPheIleGlnThrGlyHisProCysGlyProAsn	480
Qy	1509	CAATGATCTGTATAGATGGAGTTTGTATGAGTGGGATAAACAATGTTACAGACACATTT	1568
Db	481	GlnTrpValCysIleAspGlyValCysMetAsnGlyAspLysGlnCysMetAspThrPhe	500
Qy	1569	GGCAAGAAGTAGAGTTTGGCCCTTCAGAAATGTTATTCTCACCTTAAATTCAAAGACTGAT	1628
Db	501	GlyGlyGluAlaGluPheGlyProThrGluCysTyrSerTyrLeuAsnSerLysThrAsp	520
Qy	1629	GTATCTGGAACTGTGGTATAGTATCAGGATACACAGTGTGAAGCTGCAATCTG	1688
Db	521	ValSerGlyAsnCysGlyIleGlyAspSerGlyTyrThrGlnCysGluAlaAspAsnLeu	540
Qy	1689	CAGTGGGAAAAATTAATATGTAATATGAGTAAATTTTATTAACAAATTTCAAGAGGCC	1748
Db	541	GlnCysGlyLysLeuIleCysLysTyrAlaGlyLysLeuLeuGlnIleProArgAla	560
Qy	1749	ACTATTATTATGCCAACATAAGTGGACATCTCTGCATTTGCTGTGGAATTTGCGAGTGT	1808

```
Db 561 ThrilleleTyAlaAsnIleSerGlyHisLeuCysValAlaValGluPheAlaSerAsp 580
Qy 1809 CATCAGACGACCAAGAGATGTGNTAAAGATCGAACTCTTCTGGTCCAAATAAGGTT 1868
Db 581 HisGluAspSerHisLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysVal 600
Qy 1869 TGCAGGAATCAAGATGTGTAGTCTTCTCATCTTGGGTATGATGTGACTACTGACAAA 1928
Db 601 CysLysHsnGlnArgCysValSerSerSerTyrlleuGlyTyrlAspCysThrAspLys 620
Qy 1929 TGCATGATGAGGTGTATGCAATAACAAAAGCACGTCTACTGTAGTCTTCAATTTA 1988
Db 621 CysAsnHisArgGlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrl 640
Qy 1989 CTTCCAGATTCTCAGTTCATCATGATCTATGGCTGTGGTGGATATTGACAGTGGCAAT 2048
Db 641 ProProAspCysSerValGlnSerAspThrSerProGlyGlySerIleAspSerGlyAsn 660
Qy 2049 TTTCCACCTGTAGCTATACACGACGAGCTCCCTCAAGGCGCTACATTTGAGAACTTAC 2108
Db 661 PheProLeuValAlaValProAlaArgLeuProGluArgArgHisMetGluAsnValTyrl 680
Qy 2109 CATTCCAAACAAATGAGATGCCATTTCTTCTATTCTACTCTTCTTCTTATTCTGT 2168
Db 681 HisSerLysProMetArgTrpProLeuPheLeuPheIleProPhePheIlePheCys 700
Qy 2169 GTACTGATCTCTAATGTTGAAGTAAATTTCCAAAGGAAATGAGAACTGAGGAC 2228
Db 701 ValLeuIleAlaIleMetValLysValHisPheGlnArgLysLysTrpArgThrGluAsp 720
Qy 2229 TATTCAAGCGATGAGCAACTGAAAGTGAGAGTGAGAGTCAACTAAAGG 2273
Db 721 TyrSerThrAspGluGlnProGluSerGluSerGluSerProLysGly 735

RESULT 3
S47656
tMDC II protein - crab-eating macaque
C/Species: Macaca fascicularis (Crab-eating macaque)
C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C/Accession: S47656
R/Perry, A.C.F.; Barker, H.L.; Jones, R.; Hall, L.
Biochim. Biophys. Acta 1207, 134-137, 1994
A/Title: Genetic evidence for an additional member of the metalloproteinase-like, disint
C/Reference number: S47656; MUID:94318664; PMID:8043604
A/Accession: S47656
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-756 <PER>
C/Cross-references: UNIPROT:Q28483; EMBL:X77619; NID:9531477; PIDN:CAA54713.1; PID:95314
C/Superfamily: mouse meltrin alpha; disintegrin homology
P/388-473/Domain: disintegrin homology <DIS>

Alignment Scores:
Pred. No.: 1.91e-94 Length: 756
Score: 1384.50 Matches: 296
Percent Similarity: 53.85% Conservative: 124
Best Local Similarity: 37.95% Mismatches: 294
Query Match: 28.87% Indels: 66
DB: 2 Gaps: 18

US-10-054-683-18 (1-2640) x S47656 (1-756)
Qy 78 GTCTTGTCTTCAGCGGGCTCGCGGGCTGGGATGGACAGTAATTTTGATGTTA 137
Db 4 LeuLeuValLeuLeuThrGlyLeuGlyGlyMetHisAlaAspLeuAsnProHisLysThr 23
Qy 138 CCTGTGCAAAATTACAGTTCGGGAGAAATACGGTCAATA---ATAAAGGAAGCAATTGAA 194
Db 24 PheLeuGlnThrIleProGluLysIleSerSerSerAspAlaLysThrAspProGlu 43
Qy 195 TCGCAGGCATCCTACAAAATTGTAATTTGAAGGAAACCATATCTGTGAATTTAATGCAA 254
Db 195 TCGCAGGCATCCTACAAAATTGTAATTTGAAGGAAACCATATCTGTGAATTTAATGCAA 254
```

```
Db 44 HisAsnValValTyrlMetIleThrIleGluGlyLysProTyrlPheValHisLeuLysLys 63
Qy 255 AAAAATCTTTTACCCCATATTTTACAGTTTACAGTTTATAGTGGCACAGGAATTT----- 308
Db 64 GlnSerIleLeuSerSerAlaSerPheIleHisSerTyrlAspLysAsnAspIleArgHis 83
Qy 309 ATGAAACCACTTCCACCAAGATTTTCCAGAAATTTCTGCCACCTACCAAGGATTAATTTGAAGGT 368
Db 84 SerLysProLeu-----LeuValGlnMetAspCysAsnTyrlAsnGlyTyrlValAlaGly 101
Qy 369 TATCCAAATCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 428
Db 102 IleProAsnSerLeuValThrLeuSerValCysSerGlyLeuArgGlyThrMetGlnLeu 121
Qy 429 GAAATGTTAGTTATGGAATAGAACCCCTGGAGTCTTCACTTGGCTTTGAACATGTAATT 488
Db 122 LysAsnIleSerTyrlGlyIleGluProMetGluAlaValSerGlyPheIleHisLysIle 141
Qy 489 TACCAAGTAAACATTAAGAAAGCAGATGTTTCTTATATATATATATATATATATATATATATCA 548
Db 142 TyrGluGlyLysPheAlaAspThrAsnIleLeuLeuGluGluAsnAspThrTyrlSerTip 161
Qy 549 AGAGATCTGTCTTCTTAAATTAACAGCGCAGACCCAGCAAGATTTTGCAGAG----- 602
Db 162 PheAsnSerGluTyrlGlnValArgLysSerSerSerGluLysThrAspPheIleLysLeuPhe 181
Qy 603 -----TATATAGAAATGCAATGTTATAGTTGAAACAAATGTTATATATATATATATATATATAT 656
Db 182 ProArgTyrlIleGluMetHisIleValAlaAspLysAsnLeuPheAspTyrlMetGlySer 201
Qy 657 GATCAACATCTGTCGCTCAAAAGTTTTCCAGTTGATTTGATTTGATTTGATTTGATTTGATTTT 716
Db 202 AspIleAsnAlaValThrGlnLysValIleGlnIleGlyLeuValAsnThrMetLeu 221
Qy 717 GTTTCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 776
Db 222 ThrGlnLeuGlnLeuThrValIleIleSerSerIleGluIleTrpSerAsnLysAsnLys 241
Qy 777 ATTGCAACCACTGAGAGAGCTAATGATTTATTAACACATTTTAAAGATGAGAAACATCT 836
Db 242 IleSerThrThrGlyHisAlaGluTyrlValLeuLeuGluPheGluIleTrpLysLysAsp 261
Qy 837 TATCTGTTTTCATGCTCATGATGGCATTTTCTTGTGTTTACAGAGAAAGTCAAAAT 896
Db 262 HisLeuAsnPheLysProHisGlnIleAlaTyrlLeuPheValTyrlArgLysLeuProThr 281
Qy 897 TATGTTGGTCAACCTTTCAAGGAGAGTGTGATGCAAACTATGCAAGAGGTGTTGTT 956
Db 282 LeuIleGlyAlaThrPheProGlyGlnValCysAsnLysAspPheAlaAlaValAla 301
Qy 957 CTGACCCCAAGAACCACTGCTGGAAATCACTTGGAGTTATTTAGCTCAATTTATTTAGC 1016
Db 302 LeuTyrlProGluLysLeuSerLeuGluSerTyrlThrValIleIleValGlnLeuLeuGly 321
Qy 1017 CTTAGTATGGGATCACTTATGATGACATTAACAAATGCGCAGTGTCTCAGAGGTGTGTC 1076
Db 322 LeuAsnLeuGlyLeuThrTyrlAspLysThrAspThrCysHisCysSerGlyAspValCys 341
Qy 1077 ATTATGAATCCAGAACCAATTCATTCAGTGGTGTGAGATCTTTAGTAAGTCTGAGCTTC 1136
Db 342 ThrMetThrProLysAlaValTyrlSerGlyValLysAspPheSerValCysSerLeu 361
Qy 1137 GAAGACTTTGACATTTTATTTTCAAGCAGAGAGTCCAGTGTCTTCAATATCAGCTCGC 1196
Db 362 AspAspPheLysTyrlSerSerHisAsnGlyLeuThrCysLeuGlnThrAsnProLeu 381
Qy 1197 TTAGATCTTTTTCACAGCAGCAGCAGTGTGTGGTAATGCAAGCTGGAAGCAGGAG 1256
Db 382 GluMetProThrTyrlThrGlnArgArgIleCysGlyAsnGlyLeuLeuGluGlyGlyGlu 401
Qy 1257 GAGTGTGACTGTGGGACTGAACAGGATTTGTCCTTATTTGGGAAACATCTGCTGTATTT 1316
Db 402 GluCysAspCysGlyAsnLysAspAsnCysThr-----HisLysLeuCysCysAspAla 419
```





```
Db 212 GILYsValPheGlnIlePheGlyLeuIleAsnThrMetPheSerGlnLeuAsnIleThr 231
QY 735 ATTATTCCTCTTCACTGGAGCTTGGATAGTAATAAATGCAACACTGGAGAA 794
Db 232 ValMetLeuSerSerLeuGluIleThrPheSerGlnAsnIleSerThrSerGlyHis 251
QY 795 GCTAATGAGTATTACACACATTTTAAAGATGGAACATCTTATCTTGTGTTTACGTCCT 854
Db 252 AlaAspGluIleGlnArgPheLeuLeuThrPheGlnLeuLeuPheGlnArgSer 271
QY 855 CATGATGGCAATTTTACTGTTTACAGAGAAAGTCAAAATGTTGGTGCACACTTT 914
Db 272 HisAspMetThrThrLeuLeuIleThrArgAsnHisSerThrThrValGlyAlaThrThr 291
QY 915 CAAGGGAAGATGTGTGATGCAAACTATCGCAGGAGGTGTTGTTCTGCACCCAGAACATA 974
Db 292 HisGlyMetAlaCysAspProLysPheAlaThrGlyIleAlaLeuThrProLysIle 311
QY 975 AGTCTGGAATCACTTGCAGTATTATTAGCTCAATATTAGCCCTAGTATGGGATCACT 1034
Db 312 ThrValGluAlaPheSerValMetAlaGlnLeuLeuGlyIleAsnLeuGlyLeuThr 331
QY 1035 TATGATGACATTACAAATGCCAGTCTCAGAGAGTGTCTGCATATGAATCCAGAGCA 1094
Db 332 TyrAspAspIleThrAsnCysThrCysProGlyProThrCysIleMetAsnProAspAla 351
QY 1095 ATTCAATTCAGTGTGTGAAGATCTTTAGTAACCTGCAGCTTCGAGAGCTTTGCACATTT 1154
Db 352 IleArgSerHisGlyMetLysPhePheSerSerCysSerMetAspGluPheLysGlnIle 371
QY 1155 ATTTCAAAGCAGAAAGTCCCAGTGTCTTCAATCAGCTCGCTTATGATCTCTTTTCAA 1214
Db 372 ValLeuGlnProGluLeuLysCysPheGlnAspLysThrIleSerGluMetThrThrGln 391
QY 1215 CAGCAACAGTGTGTGTAATGCAAGCTGGAACGAGGAGGAGTGTGACTGTGGACT 1274
Db 392 LysSerAlaThrCysGlyAsnGlyIleLeuGluProThrGluGlnCysAspCysGlyTyr 411
QY 1275 GAACGAGTGTGCGCTTATTGGAGAAACATGCTGTGTATATTGCCACATGTAGATTAAA 1334
Db 412 ArgLysAlaCysThrPhe-----LysLysCysCysAsnProAlaAspCysThrLeuIle 429
QY 1335 GCGGTTCAAACCTGTGTAAGGACCATGCTGCGAAAC-----TGCTATTATTATGCA 1388
Db 430 GlyPheAlaGluCysGlySerGlyProCysCysAsnAsnLysThrCysThrIleTyrAla 449
QY 1389 AAGAGAGAGATGTAGTCCCTCTTGAAGAATGCGACCTCCCTGTAATATGCAATGGA 1448
Db 450 ArgGlyHisValCysArgLysSerIleAspMetCysAspPheProGluTyrCysAsnGly 469
QY 1449 TCATCTGCATCATGCCCAAGAAACCATATGTTTCAGACTGGGCATCGCTGTGGACTGAAT 1508
Db 470 ThrSerGluPheCysValProAspValLysAlaAlaAspLeuGluTyrCysSerAsnLys 489
QY 1509 CAATGGATCTGTATAGATGAGTTTGTATGAGTGGGATAAACAATGATACAGACATTT 1568
Db 490 ThrSerThrCysPheLysGlyIleCysArgAspArgAspArgGlnCysSerGlnLeuPhe 509
QY 1569 GGCAGAAAGTATAGTTTGGCCCTTCAGAAATGTTATCTCACCTTAATTCAGAGCTGAT 1628
Db 510 GlyLysPheSerLysSerAlaAsnLeuLeuCysThrGluGluValAsnPheGlnAsnAsp 529
QY 1629 GTATCTGGAAACTGTGTATAGTATAGTATCAGGATACACAGACTGTGAAGCTGCAATCTG 1688
Db 530 LysPheGlyAsnCysGly-----GlnArgCysProPhePheAspIle 543
QY 1689 CAGTGGGAAATTAATATGTAATATGTAAGTAAATTTTATACAAATTCAGAGGCC 1748
Db 544 LeuCysGlyLysIleValCysHisTrpIleHisSerGluLeuValProMetThrThrAspLeu 563
QY 1749 ACTATTATTATGCAACATAAGTGACATCTCTGCAATTCCTGTGGAATTTGCCAGTAT 1808
```

```
Db 564 AspIleGlnTyrThrTyrLeuGlyGlyHisValCysLeuSerAlaHis---AlaArgAsn 582
QY 1809 CATGCACACACCCAAAGATGTGGATAAAGATGGAACACTTCTTGTGTTCAATAAGTTT 1868
Db 583 AlaSerLysGlnLeuGlyThrTyrThrGluAspIleThrAlaCysGlyGlnGlnLysVal 602
QY 1869 TGCAAGAAATCAAGATGTGTGAGTTCCTCATCTTCTGTTTATGTTACTTACTTGACAAA 1928
Db 603 CysArgSerArgGlyCysArgTyrLeuSerGluLeuAsnIleThrLysCysThrThrAsn 622
QY 1929 TGCAATGATAGAGTGTATGCAATAACAAAGACACTGTCTGCTGTAGTGTCTTATTTA 1988
Db 623 CysGlyGlnAsnGlyIleCysAsnAspArgPheHisCysGlnCysAspProGlyTyrAla 642
QY 1989 CTTCCAGATTGC-----TCAGTTCAATCAGATCTATGGCTGTGGAGTATTTGACAGT 2042
Db 643 ProProAspCysAspProSerMetSerSer-----ProGlyGlySerIleAspAsp 659
QY 2043 GGCATTTTCCACCTGTA-----GCTATACACGCCAGACTCCTCGAAAGCGCTACATT 2096
Db 660 GlyPheThrLeuSerValAspLysSerValProLeu---PheProLysGlnArgAlaAla 678
QY 2097 GAGAACATTTACCATTCACCAACCAATGAGATGCCATTTTCTTATTATTCTTCTTTT 2156
Db 679 ProLys-----AsnAsnGlyLeuLeuIleSerPheTyrIlePheSerProLeuLeu 695
QY 2157 ATTATTTCTGTCTACTGATTGCTATTAATGCTGAAGTTTAATTTCCAAAGGAAATGCG 2216
Db 696 IleLeuIleAlaIle-----ValSerLeuLysTrpAsnLysMetLysArgPheTrp 712
QY 2217 AGA-----ACTGAGGACTATTTCAGCCGATGAGCAA 2246
Db 713 SerLysValGlyThrValSerSerArgSerIleSerGluAsp---SerSerSerAsnAsn 731
QY 2247 CCTGAAGTGAAGT 2261
Db 732 SerGlnSerGlnSer 736
RESULT 5
S18968
Cytirestin precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: I48784; S18968
R/Senftleben, A.; Wallat, S.; Lemaire, L.; Heinlein, U.A.O.
Dev. Growth Differ. 36, 49-58, 1994
A/Title: Pre and postmeiotic germ cell specific expression of TAZ83, a gene encoding a p
A/Reference number: I48784
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-823 <RES>
A/Cross-references: UNIPROT:Q62287; EMBL:X64227; NID:954264; PID:954265
C/Genetics:
A/Gene: TAZ83
C/Superfamily: mouse meltrin alpha; disintegrin homology
F:395-480/Domain: disintegrin homology <DIS>
Alignment Scores:
Pred. No.: 1,32e-87 Length: 823
Score: 1292.00 Matches: 256
Percent Similarity: 55.01% Conservative: 156
Best Local Similarity: 34.18% Mismatches: 303
Query Match: 26.94% Indels: 34
DB: 2 Gaps: 15
US-10-054-683-18 (1-2640) x S18968 (1-823)
QY 81 TTGTTTCTG---CTCAGCGGGCTCGCGGG---CTGCGGATCGACAGTAATTTTGATAGT 134
Db 5 LeuPheLeuValLeuSerTyrLeuGlyGlnValIleAlaAlaGlyLysAspValGluThr 24
QY 135 TTACTGTGCAAAATTACAGTTCCTCGGAGAAATACGGTCAATAATAAAGGAAGGA----- 188
```

Db 25 ProLeuLeuGlnIleThrValProGluLysIleAspThrAsnIleGlnAspAlaLysGlu 44  
Qy 189 ATTGAATCGCAGGATCCTACAAAATTGAATTAAGAGGAAACCATATCTACTGTGAATTTA 248  
Db 45 AlaGluThrGlnValThrTyrValValThrIleGluGlyLysAlaTyrThrLeuGlnLeu 64  
Qy 249 ATGCAAAAACCTTTTACCCCATAAATTTAGAGTTTACAGTTATAGTGGCACAGGAATT 308  
Db 65 GluLysGlnSerPheLeuHisProLeuPheGlyThrTyrLeuArgAspLysLeuGlyThr 84  
Qy 309 ATGAACCACTTGACCAAGATTTTCAGAAATTTCTGCCACTACCAAGGGTATATTGAAGGT 368  
Db 85 LeuGlnProTyrPheSerLeuValLysThrHisCysPheTyrGlnGlyHisAlaLysGlu 104  
Qy 369 TATCCAAAATCTGTGGTGTAGTGGTACATGTACTGGACTCAGGGCGCTACTACAGTTT 428  
Db 105 IleProValSerThrValThrLeuSerThrCysSerGlyLeuArgGlyLeuGlnLeu 124  
Qy 429 GAAATGTTAGTTATGAATAGAACCCCTGGAGTCTTCAGTTGGCTTTGAACATGTAATT 488  
Db 125 GluAsnIleThrTyrGlyIleGluProLeuGluSerSerAlaThrPheGluHisIleLeu 144  
Qy 489 TACCAAGTAAACATAGAAGCAGATGTTTCC-----TTATATAAT 530  
Db 145 TyrGluIleLysAsnAsnLysIleAspTyrSerProLeuLysGluAsnPheAlaAsnSer 164  
Qy 531 GAGAAGGATATTGAATCAAGAGATCTGCTCTTAAATTCACAAAGCGCAGACCCACGAA 590  
Db 165 GluGlnGluSerGlnSerTyrArgIleLeuValLysProGluLysGlySerAsnSerThr 184  
Qy 591 GATTTTGCAGATATAGAAATGCATGTTATAGTTGAAACAAATTTGATATATCATATG 650  
Db 185 LeuThrLysArgIleLeuArgIleLysIleMetAspLysAlaMetPheAspHisMet 204  
Qy 651 GGCTCTGATCAACTGTTGCTGCTCAAAAGATTTCCAGTTGATGGATTGACGAATGCT 710  
Db 205 GlySerGluValGlyValAlaThrGlnLysValValHisIlePheGlyLeuIleAsnThr 224  
Qy 711 ATTTTGTGTTTCATTATAATATTAATATTCTGCTTCATTGAGCTTTGATAGATGAA 770  
Db 225 MetPheSerGlnLeuLysMetThrValMetLeuAsnSerLeuGluIleIleTrpSerGluGln 244  
Qy 771 AATAAATCCACCACTGAGAGCTGAATAGTTATTAACACATTTTAAAGATGAAA 830  
Db 245 AspLysIleGluThrAsnGlyAspAlaAspGluValLeuGlnArgPheLeuLeuTrpLys 264  
Qy 831 ACATCTATCTGTTTACGCTCCTCATGTCGATGTCGATTTTACTGTTTACAGAGAAAG 890  
Db 265 SerLysGluIleSerGlnLysAlaGlnAspIleThrTyrLeuLeuLeuTyrLysAspHis 284  
Qy 891 TCAAAATATCTGTCGCAACCTTCAAGGCAAGATGTCATGCCAACTATGCAGGAGGT 950  
Db 285 ProAspTyrValGlyAlaThrTyrHisGlyMetAlaCysAsnProAsnPheThrAlaGly 304  
Qy 951 GTTGTCTGCACCCGCAACCAAGCTCGAATCACTTCGAGTTATTTAGCTCAATTA 1010  
Db 305 IleAlaLeuHisProLysThrLeuAlaValGluGlyPheAlaIleValLeuSerGlnLeu 324  
Qy 1011 TTGAGCCTTAGTATGGGATCACTTATGATGACATTAACAAATGCCAGTCTCAGGAGCT 1070  
Db 325 LeuGlyIleAsnLeuGlyLeuAlaTyrAspAspValTyrAsnCysPheCysProGlySer 344  
Qy 1071 GTCTGATATTAATCAGAGCAATTCATTTCAGTGGTGTGAGATCTTTAGTAACTGC 1130  
Db 345 ThrCysIleMetAsnProSerAlaIleArgSerGlnGlyIleLysValPheSerSerCys 364  
Qy 1131 AGCTTCGAAGACTTTGCACATTTTATTTCAAGCAGAGTCCAGCTGCTTTCACATCAG 1190  
Db 365 SerValAspGluPheLysGlnLeuAlaSerGlnProGluLeuAspCysLeuArgAsnThr 384  
Qy 1191 CCTCGCTTAGATCCTTTTTCACACAGCAA-----GCAGTGTGTGTTAATGCAAGCTG 1244  
Db 1244

Db 385 SerGluThrGluPheValValGlnProGlnGlySerTyrCysGlyAsnHisLeuLeu 404  
Qy 1245 GAACAGCAGGAGGAGTGTGACTGTGGGACTCAACAGCATTTGTGCCCTTATTGGAGAAACA 1304  
Db 405 GluValProGluGlnCysAspCysGlyProGluThrCysThr-----HisLysLys 422  
Qy 1305 TGCTGTGATATTGCCCATGTAGATTTAAAGCCGGTTCAAACCTGTGCTGAAGACCATGTC 1364  
Db 423 CysCysAsnProLysAspCysThrLeuIleAspAlaGlnCysGlyThrGlyProCys 442  
Qy 1365 TGGCAA-----AACTGTCTATTATGTCAAAAGAAAGATGTAGSCCTTCCTCTTGA 1418  
Db 443 CysAspLysArgThrCysThrIleAlaGluArgLysCysArgLysSerLysAsp 462  
Qy 1419 GAATGGCACTCCCTCAATATTCATATTCATCTGCATCATGCCCAAGAAACCACTAT 1478  
Db 463 GlnCysAspPheProGluPheCysAsnGlyGluThrGluGlyCysAlaProAspThrLys 482  
Qy 1479 GTTCAGACTGGGCATCCGTGTGAGCTGAATCAATGGATCTGTATAGATGGAGTTGTATG 1538  
Db 483 AlaAlaAspLeuGluProCysAsnAsnGluThrAlaTyrCysPheGlyGlyValCysArg 502  
Qy 1539 AGTGGGATAAAACAAATGTACAGACATTTTGGCAAGAAAGTGTAGCTTTGCCCTTCA 1595  
Db 503 AspProAspArgGlnCysThrAspLeuPheGlyLysTyrAlaLys---GlyProAsnTyr 521  
Qy 1596 GAATGTATTCTCACCTTAATTCAAAGACTGTATCTCGAAACTGTGTGTATTAAGTAT 1655  
Db 522 ValCysAlaGlnGluValAsnLeuGlnAsnAspLysPheGlyAsnCys----- 537  
Qy 1656 TCAGGATACACAGTGTGAAGCTGACATCTCGAGTCGCGAAATAATTAATGTAATAT 1715  
Db 538 -----HisGlyArgCysAsnTyrSerAlaIlePheCysGlyLysAlaValCysTyrTrp 555  
Qy 1716 GTAGGTAAATTTTATACAAATTCAGAGCCACTATTATTATGCAACATTAAGTGA 1775  
Db 556 AsnPheAlaGluValIleGlnThrGluLysTyrAspValGlnTyrThrTyrLeuGlyGly 575  
Qy 1776 CATCTCTGCATCTGCTGGAAATTTGCCAGTGTATCATGCACAGCCAAAGATGTGATA 1835  
Db 576 GlnValCysValSerAlaHisLeuArgSerGlnThrGlyThrArgAspAspThrTyrVal 595  
Qy 1836 AAGATGGAATCTTGTGTTCAATAAGTTTGCAGGAATCAAGATGTGTGAGTTCT 1895  
Db 596 HisAspGlyThrValCysGlySerGlyGlnValCysPheArgGlyAspCysLeuArgVal 615  
Qy 1896 TCATACTTG-----GGTTATGATTGT---ACTACTGACAAATGCAATGATAGAGTGA 1946  
Db 616 HisValLeuArgGlyThrArgGluCysGluAlaAspAspLysCysGlnGlyHisGlyIle 635  
Qy 1947 TGCAATAACAAAGAACCTGTCATCTGTAGTCTTCATATTACTCCAGATTTGCTCAGTT 2006  
Db 636 CysAsnAsnLeuAsnAsnCysGlnCysGluSerGlyPheAlaProProGluCysAspMet 655  
Qy 2007 CAATCAGATCTATGGCTGTGGGAGTATTGACAGTGGCAATTTTCCACCTGTAGCTATA 2066  
Db 656 ThrProSerSer---ProGlyGlySerMetAspAspGlyPheTrpLeuProPheAspLys 674  
Qy 2067 CAGCCAGACTCCTCGAAAGCCCTACATTGAGAACATTTACCATTCCCAACCAATGAGA 2126  
Db 675 SerThrProLeuIlePheLysArgHis-----GlyLeuLysTyrLysValLeuLeu 692  
Qy 2127 TGGCCATTTTCTTATTCATCTCTTCTTTTATTATTCTTCTGTGTACTGTTGCTTATATG 2186  
Db 693 IleSerPheTyrIleLeuLeuProPheLeuValVal-----LeuAlaPheMet 708  
Qy 2187 GTGAAAGTTAAATTTCCAAAGGAAAAATGGAGAACTGAGGACTATTTCACCGCATGACAA 2246  
Db 709 AlaValLysArgMetIleGlyLysArgLeuAlaLysGlnAsnIleSerLysAlaLeuGlu 728  
Qy 2247 CCTGAAAGTGAGGTGAACCTAAAGGG 2273  
Db 729 HisLysGluGluAlaPheAsnArgGly 737

RESULT 6

I48100

ADAM 5 protein precursor - guinea pig

C:Species: Cavia porcellus (guinea pig)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004

C:Accession: I48100

R:Wolfsberg, T.G.; Straight, P.D.; Gerena, R.L.; Huovila, A.P.; Primakoff, P.; Myles, D.

Dev. Biol. 169, 378-383, 1995

A:Title: ADAM, a widely distributed and developmentally regulated gene family encoding

A:Reference number: I48100; MUID:95269891; PMID:7750654

A:Accession: I48100

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-777 <RRS>

A:Cross-references: UNIPROT:Q60472; EMBL:U22060; NID:g965005; PIDN:AAA74918.1; PID:g9650

C:Genetics:

C:Superfamily: mouse meltrin alpha; disintegrin homology

F:395-480/Domain: disintegrin homology <DIS>

Alignment Scores:					
Pred. No.:	2.6e-87	Length:	777		
Score:	1288.00	Matches:	272		
Percent Similarity:	53.26%	Conservative:	136		
Best Local Similarity:	35.51%	Mismatches:	304		
Query Match:	26.86%	Indels:	54		
DB:	2	Gaps:	18		

US-10-054-683-18 (1-2640) x I48100 (1-777)

QY	78	GTCTGTGTTTCTGCTCAGCGGCTCGCGGGCTGGACAGTAATTTTGATAGTTTA	137
DB	4	ValLeuValLeuLeuThrGlyLeuGlyArgLeuTyrAlaGlyAsnAsnProArgLysThr	23
QY	138	CCTGTGCAAAATACAGTTCGCGAGAAATACGGTCAATA--ATAAAGAGGAAGAAATGAA	194
DB	24	PheValGlnThrThrValProGluArgIleSerSerValAspThrArgHisLeuGlu	43
QY	195	TCGCAGGCATCCTCAAAATGTAAATGAAGGAAACCATATACTGTGAATTTAAATGCAA	254
DB	44	HisAsnValAlaTyrAsnIleThrLeuLysGlySerTyrValValArgLeuLysLys	63
QY	255	AAAACTTTTACCCTAATTTTAGAGTTTACAGTTATAGTGCACAGAAATTTATGAAA	314
DB	64	GluSerPheLeuSerSerGlySerValIleTyrPheTyrAspAsnArgGlyValGlnArg	83
QY	315	-----CCACTGTGACCAAGATTTTTCAGAAATTTCTGCCACTACCAAGCGGTATATTGAAGGT	368
DB	84	SeiGlnProLeuLeuProGluMetAsp-----CysSerTyrSerGlyTyrValAlaGly	101
QY	369	TATCCAAAATCTGTGGTGTAGTGGTTAGACATGTACTGGACTCAGGCGGTACTACAGTTT	428
DB	102	PheProHisSerArgValPheAlaThrCysLeuGlyLeuArgGlyValIleGlnPhe	121
QY	429	GAAATGTGTAGTTATGGAATAGAACCCCTGGAGTCTTCAGTTGGCTTTGAACATGTAAT	488
DB	122	GluAsnValSerTyrAlaIleGluProLeuGluValLeuSerGlyPheThrHisMetIle	141
QY	489	TACCAAGTAAACATAAGAAAGCAGATGTTTTCCTTATATATATGAGAAGGATATTGAATCA	548
DB	142	TyrGluGluAsnAsnAspAsnThrHisValProLeuPheGlyLysAsnAsnSerTyrAla	161
QY	549	AGATCTGTCTTAAATTAACAAGCGCAGAGCCACAGCAAGATTTTGCA-----	599
DB	162	ArgIleHisAsnLeuGluSerGlnGlyArgSerValHisLysThrThrValSerLys	181
QY	600	-----AAGTATATAGAATGCATGTTATAGTTCGAAACAAATTTGATATCATATG	650
DB	182	LeuSerProArgTyrIleAspMetTyrIleValValAsnLysAsnLeuPheAspTyrLeu	201
QY	651	GGGTCTGATACAACTGTTGTGCTGCATCAAAAAGTTTCCAGTTGATGATGACGAATGCT	710



Db	207	ValLeuProGln-----ThrArgTyrValGluLeuPheIleValAlaAspLysGlu	223
Qy	636	TTGTATAAATCATATGGGGTCTGTACATAACCTGTTGTGCTCAAAAAGATTTTCCAGATTCATT	695
Db	224	ArgTyrAspMetMetGlyArgAsnGlnThrAlaValArgGluGluMetIleLeuLeuAla	243
Qy	696	GGATTGACGAATGCTATTTTGTTCATTTAATATTACAATATTCTGTCTTTCATTGGAG	755
Db	244	AsnTyrLeuAspSerMetTyrIleMetLeuAsnIleArgIleValLeuValGlyLeuGlu	263
Qy	756	CTTTGGATAGATGAATAAATAATTTGCACCTGGAGAGACTAATGAGTTATTACACACA	815
Db	264	IleTrpThrAsnGlyAsnLeuIleAsnIleValIleGlyAlaGlyAspValLeuGlyAsn	283
Qy	816	TTTTTAAGATGGAAAACATCTTATCTTGTGTTTAA---CGTCCTCATGATGGCAATTTTA	872
Db	284	PheValGlnTrpArgGluLysPheLeuIleThrArgArgArgHisAspSerAlaGlnLeu	303
Qy	873	CTTTGTTTACAGAGAAAAGTCAAAATATTATGTTGGTGCAACCTTTCAAGGGAAGATGTGTGAT	932
Db	304	ValLeuLysLysGlyPheGlyGlyThrAlaGlyMetAlaPheValGlyThrValCysSer	323
Qy	933	GCAAACTATGCAGAGAGTGTGTTCTGCACCCACAGAACCAATAGTCTGGAAATCACTTGCA	992
Db	324	ArgSerHisAlaGlyGlyIleAsnValPheGlyGln---IleThrValGluThrPheAla	342
Qy	993	GTTATTTTACTCAATATTATTGAGCCTTAGTATGGGATCACTTATGATGATACATTTAACAAA	1052
Db	343	SerIleValAlaHisGluLeuGlyHisAsnLeuGlyMetAsnHisAspAspGlyArgAsp	362
Qy	1053	TGCAGTGTCTCAGGAGCTGTCTGCATTATCAATTCAGAAAGCAATTCATTTCAGTGGTGTG	1112
Db	363	CysSerCysGlyAlaLysSerCysIleMetAsnSerGlyAla-----SerGlySer	379
Qy	1113	AAGATCTTTTGTAGTAATCTGCAGCTTCGAAAGACTTTGCACATTTTATTTCAAAGCAGAGTCC	1172
Db	380	ArgAsnPheSerSerCysSerAlaGluAspPheGluLysLeuThrLeuAsnLysGlyGly	399
Qy	1173	CAGTGTCTTCACAAATCAGCCTCGCTTAGATCTCTTTTTCAAAACAGCAAGCAGTGTGTGGT	1232
Db	400	AsnCysLeuLeuAsnIleProLysProAspGluAlaTyrSerAlaProSer---CysGly	418
Qy	1233	AATGCAAGCTGGAAGCAGGAGAGTGACTGTGGGAGTCAACAGGATTTGCGCCTT	1292
Db	419	AsnLysLeuValAspAlaGlyGluGluCysAspCysGlyThrProLysGluCysGluLeu	438
Qy	1293	ATTGGAGAAACATGCTGTGTGATATTGCCACATGATAGATTAAAGCCGGTTCAAACCTGCT	1352
Db	439	-----AspProCysCysGluGlySerThrCysLysLeuLysSerPheAlaGluCysAla	456
Qy	1353	GAAGGCCATGCTCGGAAAACCTGCTATTATTATGTCAAAAGAAAGATGTGTAGGCCCTTCC	1412
Db	457	TyrGlyAspCysCysLysAspCysArgPheLeuProGlyGlyThrLeuCysArgGlyLys	476
Qy	1413	TTTGAAATATGCACACTCCCTCGAATATTGCAATGGATCATCTGCATCATGCCCAAAAAC	1472
Db	477	ThrSerGluCysAspValProGluTyrCysAsnGlySerSerGlnPheCysGlnProAsp	496
Qy	1473	CACTATGTTCCAGACTGGGCATCCGTTGGACTGAATCAATGGATCTCTGTATAGATGGAGT	1532
Db	497	ValPheIleGlnAsnGlyTyrProCysGlnAsnAsnLysAlaTyrCysTyrAsnGlyMet	516
Qy	1533	TGTTATGAGTGGGATAAACAATGTACAGACACATTTTGGCAAGAGATAGAGTTTGGCCCT	1592
Db	517	CysGlnTyrTyrAspAlaGlnCysGlnValIlePheGlySerLysAlaLysAlaAlaPro	536
Qy	1593	TCAGAATGTTATTCTCACCTTAATTCAAAGACTGATGTATCTGGAAGACTGTGGTATAGT	1652
Db	537	LysAspCysPheIleGluValAsnSerLysGlyAspArgPheGlyAsnCysGlyPheSer	556
Qy	1653	GATTTCAGGATACACAGTGTGAAGCTGACAACTCTGCAGTGCAGAAATTAATATGTAA	1712

Db 557 GlyAsnGluTyrIysLysCysAlaThrGlyAsnAlaLeuCysGlyLysLeuGlnCysGlu 576

Qy 1713 TATGTAGGTAATTTTATTACAAATCCAGAGCCACTATTATTTATGTCACACATAAGT 1772

Db 577 AsnVal-----GlnGluIleProValPheGlyIleValProAlaIleIleGln 592

Qy 1773 -----GGACATCTCTGATTCGTGTGGAATTT-----GCCAGTGAATCATGCA 1814

Db 593 ThrProSerArgGlyThrLysCysTrpGlyValAspPheGlnLeuGlySerAspValPro 612

Qy 1815 GACAGCCAAAAGATGCGATAAAAGATGGAACCTCTGTGGTTCAAATAGGTTTCGAGG 1874

Db 613 AspProGly-----MetValAsnGluGlyThrLysCysGlyAlaGlyLysIleCysArg 630

Qy 1875 AATCAAAAGATGTGTGAGTCTTCATCATCTCGGTTATGATTGTACTACTGAC---AAATGC 1931

Db 631 AsnPheGlnCysValAspAlaSerValLeuAsnTyrAspCysAspValGlnLysLysCys 650

Qy 1932 AATGATAGAGGT 1943

Db 651 HisGlyHisGly 654

RESULT 9

S55060

fertilin alpha-II - crab-eating macaque

C:Species: Macaca fascicularis (crab-eating macaque)

C:Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004

C:Accession: S55060

R:Perry, A.C.F.; Gichuhi, P.M.; Jones, R.; Hall, L.

Biochem. J. 307, 843-850, 1995

A:Title: Cloning and analysis of monkey fertilin reveals novel alpha subunit isoforms.

A:Reference number: S55059; MUID:95260313; PMID:7741716

A:Accession: S55060

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-825 <PR>

A:Cross-references: UNIPROT:Q28477; EMBL:X79809; NID:9794074; PID:CAA56204.1; PID:9794074.1

C:Superfamily: mouse metrin alpha; disintegrin homology

F:443-523/Domain: disintegrin homology <DIS>

F:377/Active site: Glu #status predicted

Alignment Scores:

Pred. No.: 1.69e-68 Length: 825

Score: 1033.50 Matches: 257

Percent Similarity: 48.51% Conservative: 117

Best Local Similarity: 33.33% Mismatches: 310

Query Match: 21.55% Indels: 87

DB: 2 Gaps: 29

US-10-054-683-18 (1-2640) x S55060 (1-825)

Qy 144 CAATATACAGTTCGGAGAAATACGGTCAATAAAGGAGGATTTGAATCGCAGCA 203

Db 81 GluIleValIleProLysArgLeuMetValArgGlySerGluAspSerValGluLysAla 100

Qy 204 TCCTACAAAATTTAAATTGAAGGAAACCATATACTGTGAATTTA---ATGCAAAAAC 260

Db 101 ThrTyrLeuLeuMetGlnGlyGlnHisLeuValHisLeuLysValLysArgSer 120

Qy 261 TTTTATCCCAATAATTTAGAGTTTACAGTTATAGTGGCACAGAAATATGAACACCAT 320

Db 121 HisPheValAsnAsnPheProValTyrSerTyrHis---AsnGlyIleLeuGlyGlnGlu 139

Qy 321 GACCAAGATTTTCAGATTTCTGCCACTACCAAGGTTATTCAGGTTATCCAAATCT 380

Db 140 SerProPheLeuSerHisAspCysHisTyrIleGluGlyTyrIleGluGlyValSerGlySer 159

Qy 381 GTGGTGTATGGTTACACATGCTACTGGGCTCAGGGCGCTACTACAGTTTGAATTTAGT 440

Db 160 PheValSerValAsnThrCysAlaGlyLeuArgGlyIleLeuLysGlyLysSer 179

Qy 441 TATGGAATAGAACCCCTGGAGTCTTCAGTTGGCTTTGCAACATGTAATTTACCAAGTAAA 500







Qy	930	GATCCAAACTATGACGAGAGGTGTT--GTTCTGCACCCGACCAACCATTAAGTCTGGAAATCA	986
Db	351	SerSerGlyPheAlaAlaValGluSerPheHisGluAspMetLeuLeu-----	368
Qy	987	CTTCAGCTATTATTAGCTCAATTATTACAGCTTAGTAGTGGGATCACTTATGATGACATT	1046
Db	369	PheAlaAlaLeuMetValHisGluLeuGlyHisasnLeuGlyIleGlnHisAspHis---	387
Qy	1047	AACAAATGCCAGTGC---TCAGGAGCTGCTGTCATTATGAATCCAGAAAGCAATTCATTTC	1103
Db	388	SerAlaCysPheCysArgGluLysHisPheCysLeuMethHisGluAsnIleThrLysGlu	407
Qy	1104	AGTGGTGTGAAGATCTTTAGTAAGTCTGAGCTTCGAAGACTTTGCACATTTTATTTCAAAG	1163
Db	408	SerGly-----PheSerAsnCysSerSerAspTyrPheHisGlnPheLeuArgLys	424
Qy	1164	CAGAAGTCCCAGTCTCTTCAATCAGCTCGCTTAGATCCTTTTTTCAAAACAGCAAGCA	1223
Db	425	HisLysGlyAlaCysLeuPheAsnLysProArgProArgLysArgLysArgAspSer	444
Qy	1224	GTGTGTGGTAAATGCAAAAGCTGGAAGCAGGAGAGAGGTGTGACTGTGGGACTGAACAGGAT	1283
Db	445	AlaCysGlyAsnGlyValValGluAspThrGluGluCysAspCysGlySerAla-----	462
Qy	1284	TGTGCCCTTATTGGAGAAACATGCTGTGATATTGCCACATGTAGATTAAAGCCGGTTCA	1343
Db	463	CysHisLeu-----AspProCysCysAsp---ProThrCysThrLeuLysGluHisAla	479
Qy	1344	AACGTGCTGAAGGACCATGCTCGAAACCTGCTATTATTATGTCAAAGAAAGAAATGCTGT	1403
Db	480	GluCysSerHisGlyLeuCysCysLeuAspCysThrPheArgArgLysGlyPheLeuCys	499
Qy	1404	AGGCCTCTCTTTAAGAAGTCGACCTCCCTCGAATATTGCAATGATGATCATCTGTCATATGC	1463
Db	500	ArgProThrGlnAspGluCysAspLeuProGluTyrCysaspGlySerSerAlaGluCys	519
Qy	1464	CCAGAAACCACTATGTTCAGACTGGGGCATCCGTTGGACTGTAATCAATGGATCTGTATA	1523
Db	520	ProAlaAspSerTyrLysGlnAspGlyThrLeuCysAspArgIleHisIleTyr---CysSer	538
Qy	1524	GATGGAGTTTGTAGTGGGGATAACAATGTATACACACACATTTGGCAAGAAGTAGAG	1583
Db	539	GlyGlyGlnCysLysAsnProAspAsnGlnCysValAsnIleTyrGlyTyrProAlaArg	558
Qy	1584	TTTGGCCCTTCAGAAATGTTATCTCACCTTAATCAAAGACTGATGATCTGGAACATCTGT	1643
Db	559	SerAlaProGluAspCysTyrIleSerMetAsnThrArgGlyAspArgPheGlyAsnCys	578
Qy	1644	GGT-----ATAAGTGATTCAGGATACACACAGTGTGAAGCTGACAAATCTGCAGTGC	1694
Db	579	GlyHisProThrGluAspGlnGlnThrThrValThrCysSerAspAspAsnValPheCys	598
Qy	1695	GGAAATTAATATGTAATATGTAAGTAAATTTTATTACAAATTCCAAGGCC-----	1748
Db	599	GlyLysLeuIleCysThrGlyVal-----GlnSerLeuProArgValLysAla	614
Qy	1749	-----ACTATTATTATGCCAAATAAGTGGACATCTCTGATGCTGTGGAAATTT---	1799
Db	615	GlnHisThrValIleGlnValProHisAspAsnAspTrpCysTrpSerMetAspAlaAsp	634
Qy	1800	---GCCAGTGTATGCAGACAGCCAAAGATGTGGATAAAGATGGAACCTTCTTGTGGT	1856
Db	635	AsnIleThrAspThrProAspAsnGlyAsnValHisVal-----GlyThrSerCysAla	652
Qy	1857	TCAATAAAGTTTCAGGAATCAAGAAGATGTGTAGTTCCTTCATACTTGGGTATTGATTGT	1916
Db	653	ProAsnLysValCysThrAspTyrSerCysValHisSerIleLeuLeuTyrAspCys	672
Qy	1917	ACTACTGACAAA---TGCAAATGATAGAGTGTGATGCAATTAACAAAAAGCACTGTCACTGT	1973
Db	673	ArgProGluGluSerCysHisGlyGlyValCysAsnAsnLeuArgHisCysHisCys	692
Qy	1974	AGTGCTTCATATTTACCTCCAGATTGCTCAGTTCAATCAGATCTATCGCTGTGGGAGT	2033

Db	693	GluSerGlyPhealaProProAspCysLysAenProGlyAsn-----GlyGlySer	709
Qy	2034	AttGACAGTGGCAATTTTCCACCTGTAGCTATATACCAAGCCAGACTCCCTGAAAGCGCGTAC	2093
Db	710	ValAspSerGly-----ProProGlyMetGlnValThrAenAenSerGluSerGly	726
Qy	2094	ATTGAGAACATTATCCATTCACCAACCAATGAGATGGCCA-----TTTTTC	2138
Db	727	SerGluSerIleAlaArgGlyGlnSerLeuArgGlnAspValaAspTyrLysLeuValVal	746
Qy	2139	TTATTCAATCCTTCTTTATTATTATTTCTGTGTAAGTCTATATATGGTGAAGTTAAT	2198
Db	747	LeuLeuValProLeuPheLeuVal-----LeuLeuLeuCysSerLeuLeuThrIleSer	764
Qy	2199	TTCCAAAGGAAAAATGGAGAAT-----GAGGACTATTCAAGCGATGAG	2243
Db	765	TyrLeuCysSerGluValGlnThrAlaValAlaGluValGluGluSerThrGluThr	784
Qy	2244	CAACCTGAAGTGAG	2258
Db	785	ThrLeuGluSerGlu	789

RESULT 11

S71949

metalloproteinase 12 (EC 3.4.24.-) precursor - human

N:Alternate names: disintegrin 12; myeloma cell metalloproteinase MCMP

C:Species: Homo sapiens (man)

C:Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 09-Jul-2004

C:Accession: S71949; PC2464

R:McKie, N.; Dallas, D.U.; Edwards, T.; Apperley, J.P.; Ruseell, R.G.G.; Crouch Biochem. Biophys. Res. Commun. 230, 335-339, 1997

A:Title: Cloning of a novel membrane-linked metalloproteinase from human myeloma

A:Reference number: S71949; MUID:96404892; PMID:8809033

A:Accession: S71949

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-660 <MCK>

A:Cross-references: UNIPROT:Q13443

A:Experimental source: myeloma cells

R:McKie, N.; Edwards, T.; Dallas, D.J.; Houghton, A.; Stringer, B.; Graham, R.; Biochem. Biophys. Res. Commun. 230, 335-339, 1997

A:Title: Expression of members of a novel membrane linked metalloproteinase fam

A:Reference number: PC2463; MUID:97168971; PMID:9016778

A:Accession: PC2464

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-337 <MCF>

A:Experimental source: articular chondrocyte

C:Function:

A:Description: involved in cell/cell and cell/matrix interaction in a variety o

A:Note: membrane-bound; belongs to repolysin family of metalloproteinases

C:Superfamily: mouse matriin alpha; disintegrin homology

C:Keywords: hydrolase; metalloproteinase; zinc

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-660/Product: pre-metalloproteinase 12 #status predicted <MAT>

F:54-59/Region: autoinhibitory

P:295-378/Domain: disintegrin homology <DIS>

P:574-598/Region: epidermal growth factor-like

F:622-642/Domain: transmembrane #status predicted <TM>

F:572,229,233,239/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited)

F:229,233,239/Binding site: zinc, catalytic (His) (active) #status predicted

P:230/Active site: Glu #status predicted

Alignment Scores:		
Pred. No. :	3.05e-67	Length: 660
Score:	1016.50	Matches: 216
Percent Similarity:	52.17%	Conservative: 109
Best Local Similarity:	34.67%	Mismatches: 207
Query Match:	21.19%	Indels: 91
DB:	2	Gaps: 16

US-10-054-683-18 (1-2640) x S71949 (1-660)

QY 384 GTGATGTTAGCACATGCTACTGCGGCGCTACTACAGTTTGAATTTAGTTAT 443  
Db : : : : :  
10 LeuLeuAlaThrValLeuAspLeuArgGlyLeuLeuHisLeuGluAsnAlaSerTyr 29  
QY 444 GGAATAGAACCCCTGGAGTCTTACGTGGCTTTGAAACATGTAATTTACCAAGTA----- 497  
Db : : : : :  
30 GlyIleGluProLeuGlnMetSerHisPheGluHisIleIleTyrArgMetAspAsp 49  
QY 498 -----AAACATAAGAACAGCAGATGTTCTTATATATAGAGAGGATTTGAATCA 548  
Db : : : : :  
50 ValTyrLysGluProLeuLysCysGlyValSer-----AsnLysAspIleGluLys 66  
QY 549 AGA-----GATCTGCTCTTTAAATTA 569  
Db : : : : :  
67 GluThrAlaLysAspGluGluProProSerMetThrGlnLeuLeuArgArg 86  
QY 570 CAAAGCCGAGCCACACAGCAGATTTTGCAAGATATATAGAAATGATTTAGTTGAA 629  
Db : : : : :  
87 ArgAlaValLeuProGln-----ThrArgTyrValGluLeuPheIleValValAsp 103  
QY 630 AAACAAATTGTATATATATGCGGTCTGATACAACTGTTGCGCTCAAAAAGTTTCCAG 689  
Db : : : : :  
104 LysGluArgTyrAspMetGlyArgAsnGlnThrAlaValArgGluGluMetIleLeu 123  
QY 690 TTGATTGGATTGACGAATGCTATTGTTTTCATTATTAATACAAATTTCTCTCTCA 749  
Db : : : : :  
124 LeuAlaAsnTyrLeuAspSerMetTyrIleMetLeuAsnIleArgIleValLeuValGly 143  
QY 750 TTGGAGCTTTGGATAGATGAAATAAATTCACACCTGGAGAACCTAATGATTTATTA 809  
Db : : : : :  
144 LeuGluIleTrpThrAsnGlyAsnLeuIleAsnIleValGlyGlyAlaGlyAspValLeu 163  
QY 810 CACACATTTTAAAGTGAACATCTTATCTCTTTTA---CGTCCTCATGTGCGCA 866  
Db : : : : :  
164 GlyAsnPheValGlnTrpArgGlyPheLeuIleThrArgArgGHisAspSerAla 183  
QY 867 TTTTACTGTTTACAGAGAAAGTCAAAATATGTTGTCGCACTTTCAAGGGAAGATG 926  
Db : : : : :  
184 GlnLeuValLeuLysGlyPheGlyGlyThrAlaGlyMetAlaPheValGlyThrVal 203  
QY 927 TGTGATCAAACTATGAGGAGGTGTTGTCGACCCAGAACCAATGCTGGATCA 986  
Db : : : : :  
204 CysSerArgSerHisAlaGlyGlyIleAsnValPheGlyGln---IleThrValGluThr 222  
QY 987 CTGTCAGTTATTTAGCTCAATATTAGCCTTATGAGCTTATGGGATCACTTATGATGACATT 1046  
Db : : : : :  
223 PheAlaSerIleValAlaHisGluLeuGlyHisAsnLeuGlyMetAsnHisAspAspGly 242  
QY 1047 AACAAATGCCAGTCTCAGGAGCTGTGTCATTATGAATCCAGAACCAATTTCAATTCAGT 1106  
Db : : : : :  
243 ArgAspCysSerCysGlyAlaLysSerCysIleMetAsnSerGlyAla-----Ser 259  
QY 1107 GGTGTGAAGATCTTTAGTAAGCTGAGCTTGAAGACTTTGACATTTTATTTCAAGCAG 1166  
Db : : : : :  
260 GlySerArgAsnPheSerSerCysSerAlaGluAspPheGluLysLeuThrLeuAsnLys 279  
QY 1167 AAGTCCAGTCTTCAACATCACCTCGCTTACATCTTTTTCACACGACGAGTG 1226  
Db : : : : :  
280 GlyCysLeuLeuLeuAsnIleProLysProAspGluAlaTyrSerAlaProSer--- 298  
QY 1227 TGTGGTAAATGCAAGCTGAGGAGAGAGTGTGACTGTGGAGCTGGAACAGGATGT 1286  
Db : : : : :  
299 CysGlyAsnLysLeuValAspAlaGlyGluCysAspCysGlyThrProLysGluCys 318  
QY 1287 GCCCTTATTTGAGAAACATGCTGTGATTTGCCACATGTGATTTAAAGCCGGTTCAAC 1346  
Db : : : : :  
319 GluLeu-----AspProCysCysGluGlySerThrCysLysLeuLysSerPheAlaGlu 336  
QY 1347 TGTGCTGAAGCACCATCTCGAAACTGTCTATTTATGTCARAAGAAAGATGTGAGG 1406  
Db : : : : :  
337 CysAlaTyrGlyAspCysCysLysAspCysArgPheLeuProGlyGlyThrLeuCysArg 356

QY 1407 CCTTCCTTTGAGAATGCGACCTCCCTGAATATTTGCAATGATCATCTGCAATCATGCCCCA 1466  
Db : : : : :  
357 GlyLysThrSerGluCysAspValProGluTyrCysAsnGlySerGlnPheCysGln 376  
QY 1467 GAAACCACTATGTTGAGACTGGGCATCCGTGTGGACTGAATCAATGGATCTGTATAGAT 1526  
Db : : : : :  
377 ProAspValPheIleGlnAsnGlyTyrProCysGlnAsnAsnLysAlaTyrCysTyrAsn 396  
QY 1527 GGAAGTTTGTAGTGGGATAAACAATGTACACACATTTGGCAAGAGATGAGCTTT 1586  
Db : : : : :  
397 GlyMetCysGlnTyrTyrAspAlaGlnCysGlnValIlePheGlySerLysAlaLysAla 416  
QY 1587 GGCCTTCAGAATGTTATCTCACCTTAATTAACAAAGACTGATGATTCGAAAAGTGGT 1646  
Db : : : : :  
417 AlaProLysAspCysPheIleGluValAsnSerLysGlyAspArgPheGlyAsnCysGly 436  
QY 1647 ATAAGTGTGATTCAGATACACACAGTGTGAA----- 1676  
Db : : : : :  
437 PheSerGlyAsnGluTyrLysLysCysAlaThrGlyLeuSerLeuLysPheHisAlaPro 456  
QY 1676 ----- 1676  
Db : : : : :  
457 PheLeuSerThrMetLeuGlnGluAlaValArgGlnThrGlyThrTyrLeuGlyGlySer 476  
QY 1677 -----GCTGACAAATCTGCAATGCGGA 1697  
Db : : : : :  
477 ValCysCysMetLysSerAspCysArgIleValThrLeuValLysAsnAlaLeuCysGly 496  
QY 1698 AAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1757  
Db : : : : :  
497 LysLeuGlnCysGluAsnVal-----GlnGluIleProValPheGlyIleVal 512  
QY 1758 TATGCCCAACATAAGT-----GGACATCTCTGCAATGCTGTGGAATTT 1799  
Db : : : : :  
513 ProAlaIleIleGlnThrProSerArgGlyThrLysCysTrpGlyValAspPheGlnLeu 532  
QY 1800 GCCAGTATCATCAGACACGCCAAAAGATGTTGATGATGATGATGATGATGATGATGAT 1859  
Db : : : : :  
533 GlySerAspValProAspProGly-----MetValAsnGluGlyThrLysCysGlyAla 550  
QY 1860 AATAAGTTTTCAGGATCAAGATGTTGATGATGATGATGATGATGATGATGATGATGAT 1919  
Db : : : : :  
551 GlyLysIleCysArgAsnPheGlnCysValAspAlaSerValLeuAsnTyrAspCysAsp 570  
QY 1920 ACTGAC---AAATGCAATGATGAGGTGTATGCAATAACAAAGACACTGCTGATGAT 1976  
Db : : : : :  
571 ValGlnLysLysCysHisGlyHisGlyValCysAsnSerAsnLysAsnCysHisCysGlu 590  
QY 1977 GCTTCATATTTACTCAGATGCTCAGATGCTCAGATGCTCAGATGCTCAGATGCTCAGAT 2036  
Db : : : : :  
591 AsnGlyTrpAlaProAsnCysGluThrLysGly-----TyrGlyGlySerVal 607  
QY 2037 GACAGTGGC 2045  
Db : : : : :  
608 AspSerGly 610  
RESULT 12  
S60257  
meltrin alpha - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004  
C;Accession: S60257  
R;Yagami-Hironaka, T.; Sato, T.; Kurisaki, T.; Kamijo, K.; Nabeshima, Y.I.; Fujisawa-Seh-  
Nature 377, 652-656, 1995  
A;Title: A metalloprotease-disintegrin participating in myoblast fusion.  
A;Reference number: S60257; MUID:96026308; PMID:7566181  
A;Accession: S60257  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-903 <VAG>  
A;Cross-references: UNIPROT:Q61824; EMBL:D50411; NID:gl054586; PIDN:BAA0891.1; PID:gl05-  
C;Superfamily: mouse meltrin alpha; disintegrin homology  
F;421-503/Domain: disintegrin homology <DIS>





Qy	714	TTTGTTTCATTAAATATTACAATATTCTGTCTTTCATTGGAGCTTTGGATAGATGAAAT	773
Db	233	TyrGlnGluLeuSerPheArgValValLeuValGlyLeuGluIleTrp--AsnLysAsp	251
Qy	774	AAAATTGCACCACTGGAGAGCTAATAGCTATTATACACATTTTAAAGATGAAAAACA	833
Db	252	LysPheTyrIleSerArgTfYrAlaAsnValThrLeuGluAsnPheLeuSerTrpArgGlu	271
Qy	834	TCATTATCTGTT--TTACGTCCTCATGATGTGGCATTTTACTTTTACAGAGAAAG	890
Db	272	GlnAsnLeuGlnGlyGlnHisProHisAspAsnValGlnLeuIleThr-----Gly	288
Qy	891	TCAAAATATTGTTGGTGCACCTTT-----CAAGGGAAGATGTGTGATGCAAAAC	938
Db	289	ValAspPheIleGlySerThrValGlyLeuAlaLysValSerAlaLeuCysSerArgHis	308
Qy	939	TATGCAGGAGGTGTTGTTCTGCACCCAGAACATTAAGTCTGGNAATCACTTGCAGTTATT	998
Db	309	-----SerGlyAlaValAsnGlnAspHisSerLysAsnSerIleGlyValAlaSerThr	326
Qy	999	TTAGTCAATATTATGGACCTTAGTATGGGATCACTTATGAT--GACATTAACAATATGC	1055
Db	327	MetAlaHisGluLeuGlyHisAsnLeuGlyMetSerHisAspGluAspIleProGlyCys	346
Qy	1056	CAGTGC-----TCAGGAGCTGTCTGCATTTATGAATCCAGAAAGCAATTCATTTC	1103
Db	347	TyrCysProGluProArgGluGlyGlyCysIleMetThr---GluSerIleGlySer	365
Qy	1104	AGTGGTGTGAAGATCTTTAGTAATCTGCAGCTTCGAAGACTTTCACATTTTATTTCAAAG	1163
Db	366	LysPheProArgIlePheSerArgCysSerLysIleAspLeuGluSerPheValThrLys	385
Qy	1164	CAGAACTCCAGTGTCTTCAATCAGCTTCGCTAGATCCTTTTTCAAACAGCAAGCA	1223
Db	386	ProGlnThrGlyCysLeuThrAsnValProAspValAsnArgPheVal---GlyGlyPro	404
Qy	1224	GTGTGTGGTAAATCGAAAGCTGGAAGCAGGAGAGAGTGTCACTGTGGAGCTGAACAGGAT	1283
Db	405	ValCysGlyAsnLeuPheValGluHisGlyGluGlnCysAspCysGlyThrProGlnAsp	424
Qy	1284	TGTGCCCTTATGGAGAAACATGCTGTGATATTGTCACATGTAGATTAAAGCGCGTTCA	1343
Db	425	CysGln-----AsnProCysCysAsnAlaThrThrCysGlnLeuValLysGlyAla	441
Qy	1344	AACGTGCTGAAGGACCATGCTCGGAACACTGCTATTATGTCTCAAAGAAAGATGTGT	1403
Db	442	GluCysAlaSerGlyThrCysCysHisGluCysLysValLysProAlaGlyGluValCys	461
Qy	1404	AGGCCTTCCTTTGAAGAATCGACCTCCCTCGAATATTGCAATGGATCATCTGCATCATGC	1463
Db	462	ArgLeuSerLysAspLysCysAspLeuGluGluPheCysAspGlyArgLysProThrCys	481
Qy	1464	CCAGAAAACCATATGTTTCAGACTGGGCATCCGTGTGGACTGAATCAATGGATCTGTATA	1523
Db	482	ProGluAspAlaPheGlnAsnGlyThrProCys-----ProGlyGlyTyrCysPhe	499
Qy	1524	GATGGAGTTTGTATAGTGGGGATAACATGTACAGACACATTTGGCAAGAAGTAGAG	1583
Db	500	AspGlySerCysProThrLeuAlaGlnGlnCysArgAspLeuTrpGlyProGlyAlaArg	519
Qy	1584	TTTGGCCCTTCAGAAATGTTATTCTCACCTTAAATTCAAAGACTGATGTCTGGAAACCTGT	1643
Db	520	ValAlaAlaAspSerCysTyrThrPhe-----SerIleProProGlyCys	534
Qy	1644	---GGTATATAGTATCAGGATACACACAGTGTGAAGCTGACAATCTGCAGTGGGAAAA	1700
Db	535	AsnGlyArgMetTyrSerGlyArgIleAsn-----ArgCysGlyAla	548
Qy	1701	TTAATATGTAAATATGTAGTAAATTTTATTACAATTCACAGAGCCACTATTTAT	1760
Db	549	LeuTyrCysGluGlyGlnLys-----ProLeuGluArgSerPheCysThrPhe	565



Search completed: January 10, 2005, 22:07:12  
Job time : 148.5 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 10, 2005, 21:24:40 ; Search time 491.5 Seconds  
(without alignments)  
6181.034 Million cell updates/sec

Title: US-10-054-683-18

Perfect score: 4796

Sequence: 1 catctgcactccaactgc.....actgttacatttcaaaaaa 2640

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 3650362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DBV=xlp  
-O=/cgn2 1/USPTO.spool\_p/US10054683/runat 10012005 174510 15496/app query fasta\_1.2823  
-DB=UniProt\_02 -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10054683 @CGN 1 1 518 @runat 10012005 174510 15496 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

UniProt\_02:\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3969.5	82.8	735	1 AD02 HUMAN	Q9965 homo sapien
2	3690.5	76.9	735	1 AD02 MACFA	Q28478 macaca fasc
3	3065.5	63.9	579	2 Q6P2G0	Q6p290 homo sapien
4	3065.5	63.9	579	2 AAH64547	Aah64547 homo sapi
5	2820	58.8	735	2 Q866A8	Q866a8 sus scrofa
6	2723	56.8	745	1 AD02 BOVIN	O77780 bos taurus
7	2685	56.0	751	1 AD02 RABIT	Q28660 oryctolagus
8	2479.5	51.7	737	1 AD02 RAT	Q63202 rattus norv
9	2453.5	51.2	735	1 AD02 MOUSE	Q60718 mus musculus
10	2326	48.5	735	1 AD02 CAVPO	Q60411 cavia porce
11	1586.5	33.1	746	1 AD18 MACFA	Q95194 macaca fasc
12	1582.5	33.0	739	1 AD18 HUMAN	Q9y3q7 homo sapien
13	1524.5	31.8	787	2 Q8TC27	Q8tc27 homo sapien
14	1524.5	31.8	787	2 QAQ90909	QaQ90909 homo sapi
15	1519.5	31.7	787	2 Q8TC42	Q8tc42 homo sapien
16	1493.5	31.1	715	2 Q6UXJ9	Q6uxj9 homo sapien

17	1493.5	31.1	715	2 AAQ88687	AaQ88687 homo sapi
18	1437.5	30.0	750	2 Q6P901	Q6p901 mus musculus
19	1437.5	30.0	750	2 AAH60983	Aah60983 mus muscu
20	1433	29.9	750	2 Q8K410	Q8k410 mus musculu
21	1384.5	28.9	756	2 Q28483	Q28483 macaca fasc
22	1373	28.6	730	2 Q80Y08	Q80y08 mus musculu
23	1367.5	28.5	677	2 Q8BJ80	Q8bj80 mus musculu
24	1366.5	28.5	719	1 AD18 MOUSE	Q9r157 mus musculu
25	1365.5	28.5	736	2 Q28482	Q28482 macaca fasc
26	1293	27.0	822	2 Q62287	Q62287 mus musculu
27	1292	26.9	740	2 P70534	P70534 rattus norv
28	1288	26.9	777	2 Q60472	Q60472 cavia porce
29	1237	25.8	771	2 Q8CDV5	Q8cdv5 mus musculu
30	1237	25.8	771	2 Q9D4F0	Q9d4f0 mus musculu
31	1217	25.4	819	1 AD09 HUMAN	Q13443 homo sapien
32	1206	25.1	751	2 Q60816	Q60816 mus musculu
33	1174	24.5	873	2 Q42595	Q42595 xenopus lae
34	1167.5	24.3	845	1 AD09 MOUSE	Q61072 mus musculu
35	1076	22.4	919	2 Q28659	Q28659 oryctolagus
36	1047.5	21.8	792	2 O19061	O19061 saguinus oe
37	1045	21.8	812	2 O77779	O77779 bos taurus
38	1033.5	21.5	825	2 Q28477	Q28477 macaca fasc
39	1031.5	21.5	791	1 AD1A MOUSE	Q60813 mus musculu
40	1026.5	21.4	789	1 AD01 RAT	P70505 rattus norv
41	1022.5	21.3	905	2 Q28476	Q28476 macaca fasc
42	1019.5	21.3	825	2 O46652	O46652 papio anubi
43	1016	21.2	821	2 O19060	O19060 saguinus oe
44	1014.5	21.2	838	2 O19056	O19056 papio anubi
45	1007	21.0	804	2 Q60410	Q60410 cavia porce

ALIGNMENTS

RESULT 1

ID	AD02 HUMAN	STANDARD;	PRT;	735 AA.
AC	Q99965; P78326; Q9UQ88;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)			
DE	(Fertilin beta subunit) (PH-30) (PH30).			
GN	Names=ADAM2; Synonyms=FTNB;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
[1]				
RN	SEQUENCE FROM N.A. (ISOFORM 1).			
RP	TISSUE=Testis;			
RX	MEDLINE=97193554; PubMed=9041139;			
RA	Vidaeus C.M., von Kap-Herr C., Golden W.L., Eddy R.L., Shows T.B.,			
RA	Herr J.C.;			
RT	"Human fertilin beta: identification, characterization, and			
RT	chromosomal mapping of an ADAM gene family member.";			
RL	Mol. Reprod. Dev. 46:363-369(1997).			
[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Testis;			
RX	MEDLINE=96235488; PubMed=8702389;			
RA	Gupta S.K., Alves K., Palladino L.O., Mark G.E., Hollis G.F.;			
RT	"Molecular cloning of the human fertilin beta subunit.";			
RL	Biochem. Biophys. Res. Commun. 224:318-326(1996).			
[3]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Testis;			
RX	MEDLINE=9724507; PubMed=9070941;			
RA	Burkin H.R., Burkin D.J., Davey P.M., Griffin D.K., Affara N.A.;			
RT	"Mapping, sequence, and expression analysis of the human fertilin beta			
RT	gene (FTNB).";			
RL	Genomics 40:190-192(1997).			
[4]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).			

RC TISSUE=Testis;  
RA Hall L., Frayne J.;  
RT "Nucleotide sequence of the human fertilin beta transcript.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Testis;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schenker C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T.I., Max S.I., Wang J., Haieh F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Boak S.A., McWeney P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Touchman A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Mallat D.E.,  
RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
CC -1- FUNCTION: Sperm surface membrane protein that may be involved in  
sperm-egg plasma membrane adhesion and fusion during  
fertilization. Could have a direct role in sperm-zona binding or  
migration of sperm from the uterus into the oviduct. Interactions  
with egg membrane could be mediated via binding between its  
disintegrin-like domain to one or more integrins receptors on the  
egg. This is a non catalytic metalloprotease-like protein.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC -1- Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q99965-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q99965-2; Sequence=VSP\_005471;  
CC Note=No experimental confirmation available;  
CC -1- TISSUE SPECIFICITY: Expressed specifically in spermatogenic cells  
in the seminiferous cells. Not detected in fetal tissues.  
CC -1- DOMAIN: A tripeptide motif (FEE) within disintegrin-like domain  
could be involved in the binding to egg integrin receptor and thus  
could mediate sperm/egg binding.  
CC -1- PTM: The prodomain and the metalloprotease domain are cleaved  
during the epididymal maturation of the spermatozoa.  
CC -1- MISCELLANEOUS: In mammals, exists as a heterodimer composed of an  
alpha and beta subunits. In human, fertilin alpha is a pseudogene.  
CC -1- SIMILARITY: Belongs to peptidase family M12B.  
CC -1- SIMILARITY: Contains 1 disintegrin domain.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC ENBL: U52370; AAC51110.1; --  
CC ENBL: U38805; AAD04206.1; --  
CC ENBL: X99374; CAA67753.1; --  
CC ENBL: AJ133005; CAB40813.1; --  
CC ENBL: BC034957; AAH34957.1; --  
CC PIR: JC4861; JC4861.  
CC HSSP: P30403; IN4Y.  
CC MEROPS: M12.950; --  
CC Genew; HGNC:198; ADAM2.

DR MIM; 601533; --  
DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
DR GO; GO:0005178; F:integrin binding; TAS.  
DR GO; GO:0008337; F:metallopeptidase activity; TAS.  
DR GO; GO:0007342; P:fusion of sperm to egg plasma membrane; TAS.  
DR InterPro; IPR006586; ADAM cysteine.  
DR InterPro; IPR001762; Disintegrin.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR001590; Peptidase M12B.  
DR InterPro; IPR002870; Peptidase M12B\_N.  
DR InterPro; IPR006025; Pept M\_Zn\_BS.  
DR Pfam; PF00200; Disintegrin; 1.  
DR Pfam; PF01562; Pep M12B proper; 1.  
DR Pfam; PF01421; Repolysin; 1.  
DR PRINTS; PR00289; DISINTEGRIN.  
DR PRODOM; PD000664; Disintegrin; 1.  
DR SMART; SM00608; ACR; 1.  
DR SMART; SM00050; DISIN; 1.  
DR PROSITE; PS00215; ADAM\_MBP; 1.  
DR PROSITE; PS00427; DISINTEGRIN\_1; 1.  
DR PROSITE; PS00214; DISINTEGRIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; FALSE NEG.  
DR PROSITE; PS01186; EGF\_2; FALSE NEG.  
KW Alternative splicing; Cell adhesion; EGF-like domain; Glycoprotein;  
KW Signal; Transmembrane.  
FT SIGNAL 1 16 Potential.  
FT PROPEP 17 174  
FT CHAIN 175 735 ADAM 2.  
FT DOMAIN 175 686 Extracellular (Potential).  
FT TRANSMEM 687 707 Potential.  
FT DOMAIN 708 735 Cytoplasmic (Potential).  
FT DOMAIN 175 382 Metalloprotease-like.  
FT DOMAIN 383 475 Disintegrin-like.  
FT DOMAIN 477 606 Cys-rich.  
FT DOMAIN 612 645 EGF-like.  
FT DISULFID 287 370 By similarity.  
FT DISULFID 329 354 By similarity.  
FT DISULFID 445 458 Potential.  
FT DISULFID 616 627 By similarity.  
FT DISULFID 621 633 By similarity.  
FT DISULFID 635 644 By similarity.  
FT CARBOHYD 122 122 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 220 220 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 353 353 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 459 459 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 566 566 N-linked (GLCNAC. .) (Potential).  
FT VARSPPLIC 172 190 Missing (in isoform 2).  
FT FTID=VSP\_005471.  
FT CONFLICT 3 3 Missing (in Ref. 2).  
FT CONFLICT 21 21 D -> H (in Ref. 3).  
FT CONFLICT 99 99 E -> D (in Ref. 3).  
FT CONFLICT 106 106 V -> G (in Ref. 3).  
FT CONFLICT 170 170 V -> A (in Ref. 2).  
FT CONFLICT 288 288 D -> H (in Ref. 1).  
FT CONFLICT 321 321 I -> T (in Ref. 1).  
FT CONFLICT 388 388 G -> S (in Ref. 3).  
FT CONFLICT 396 398 EEC -> DEF (in Ref. 3).  
FT CONFLICT 501 501 G -> S (in Ref. 3).  
FT CONFLICT 529 529 D -> Y (in Ref. 3).  
FT CONFLICT 579 579 S -> G (in Ref. 3).  
FT CONFLICT 588 588 W -> L (in Ref. 3).  
FT CONFLICT 603 603 N -> D (in Ref. 3).  
FT CONFLICT 629 630 NK -> KQ (in Ref. 3).  
FT CONFLICT 638 638 S -> F (in Ref. 3).  
SQ SEQUENCE 735 AA; 82456 MW; 92867B5340BEE01F CRC64;  
Alignment Scores:  
Pred. No.: 1,18e-291 Length: 735  
Score: 3969.50 Matches: 733  
Percent Similarity: 99.73% Conservative: 0  
Best Local Similarity: 99.73% Mismatches: 1  
Query Match: 82.77% Indels: 1  
DB: 1

US-10-054-683-18 (1-2640) x AD02\_HUMAN (1-735)

QY 72 ATGTGG---GTCCTGTTCTGCTCAGCGGGCTCGGGGGCTGCGGATGGACAGTAATTTT 128  
DB 1 MetTrpArgValLeuPheLeuLeuSerGlyLeuGlyLeuArgMetAspSerAsnPro 20  
QY 129 GATGATTTACCTGTGCAAAATTACAGTTCCGAGAAAATACGGTCAATAATAAGGAAGGA 198  
DB 21 AspSerLeuProValGlnIleThrValProGluLysIleArgSerIleIleLysGluGly 40  
QY 189 AFTGAATCCGAGCATCCTACAAATTTGAATTTGAAGGGAACCATATACTGTGAATTTA 248  
DB 41 IleGluSerGlnAlaSerTyrLysIleValIleGluGlyProTyrThrValAsnLeu 60  
QY 249 ATCAAAAAAATTTTACCCCATAAATTTAGAGTTTACAGTTTATAGTGGCACAGGAAT 308  
DB 61 MetGlnLysAsnPheLeuProHisAsnPheArgValTyrSerTyrSerGlyThrGlyIle 80  
QY 309 ATGAAACCACTTGACCAAGATTTTCAGAAATTTCTGCCACTACCAAGGTATATTGAAGGT 368  
DB 81 MetLysProLeuAspGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGly 100  
QY 369 TATCCAAAATCTGTGTGATGTTAGCACATGTACTGGACTCAGGGCGGTACTACAGTTT 428  
DB 101 TyrProLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPhe 120  
QY 429 GAAAAATGTTAGTTATGGAATAGAACCCCTCGAGTCTTCAGTTGGCTTTGAACATGTAAT 488  
DB 121 GluAsnValSerTyrGlyIleGluProLeuGluSerValGlyPheGluHisValIle 140  
QY 489 TACCAAGTAAACATAAGAACAGATGTTTCCTTATATATAGAGAGATATTGAATCA 548  
DB 141 TyrGlnValLysHisLysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSer 160  
QY 549 AGAGATCTGCTCTTAAATTTACAAAGCGGACGACGACCAAGATTTTGCAAAATATATA 608  
DB 161 ArgAspLeuSerPheLysLeuGlnSerValGluProGlnGlnAspPheAlaLysTyrIle 180  
QY 609 GAAATGATCTTATAGTTGAAACCAATTTCTATAATCATATGGTCTGTATACAACTGTT 668  
DB 181 GluMetHisValIleValGluLysGlnLeuTyrAsnHisMetGlySerAspThrThrVal 200  
QY 669 GTCGCTCAAAAGTTTTCCAGTTGATTGGAATGACGAATCTATTTTGTTCATTTAAT 728  
DB 201 ValAlaGlnLysValPheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsn 220  
QY 729 ATTACAAATATTCTGCTCTTCAATGGAGCTTTGGATAGATGAAATATAAATTTGCAACCACT 788  
DB 221 IleThrIleLeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThrThr 240  
QY 789 GGAGAGCTAATGAGTTATTACACATTTTAAAGATGGAAACATCTTATCTGTTTAA 848  
DB 241 GlyGluAlaAsnGluLeuLeuHisThrPheLeuArgTrpLysThrSerTyrLeuValLeu 260  
QY 849 CGTCCTCATGATGGCATTTTACTTTGTTTACAGAAAAGTCAAAATATTGTTGTGCA 908  
DB 261 ArgProHisAspValAlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAla 280  
QY 909 ACCTTTCAAGGAAGATGTGTGATGCAAACTATGCGAGGAGGTGTTGTCGACCCGAGA 968  
DB 281 ThrPheGlnGlyLysMetCysAspAlaAsnTyrAlaGlyValValLeuHisProArg 300  
QY 969 ACCATAGCTGGAATCAGTTGAGTTATTTAGCTCAATATTAGCCCTTAGTATGGGG 1028  
DB 301 ThrIleSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGly 320  
QY 1029 ATCATTATGATGACATTAAACAAATGCGAGTGTCTCAGAGCTGTCTGCATTTATGAATCCA 1088  
DB 321 IleThrTyrAspAspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnPro 340  
QY 1089 GAAGCAATTCATTTCAGTGTGTGAAGATCTTTAGTAACTTCGACGTTTCGAAGACTTTGCA 1148  
DB

DB 341 GluAlaIleHisPheSerGlyValLysIlePheSerAsnCysSerPheGluAspPheAla 360  
QY 1149 CATTTTATTTCAAAGCAGAAAGTCCCAAGTGTCTTTCACAATCAGCTCGCTTAGACTCTTTT 1208  
DB 361 HisPheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhe 380  
QY 1209 TTCAAAACAGAACAGCATGTGTGTAATGCMAAGCTGMAAGCAGGAGAGAGTGTGACTGT 1268  
DB 381 PheLysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCys 400  
QY 1269 GGGACTGAACAGGATTTGCCCTTATTGGAGAAACATGCTGTGCATATTGGCCACATGTAGA 1328  
DB 401 GlyThrGluGlnAspCysAlaLeuIleGlyLeuThrCysCysAspIleAlaThrCysArg 420  
QY 1329 TTTTAAAGCCGGTTTCAAACTGTGCTGAAGGACCACTATGCTGCGAAAACATGCTCTATTATGTCA 1388  
DB 421 PheLysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSer 440  
QY 1389 AAGAAGAATGTGTAGGCTTCCCTTTGAAGATGGACCTCCCTCGAATATTGCAATGGA 1448  
DB 441 LysGluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGly 460  
QY 1449 TCATCTGCATCATGCCCCAGAAAACCACTATGTTTCAGACTGGGCATCCGCTGTGACTGAAT 1508  
DB 461 SerSerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsn 480  
QY 1509 CAATGATCTGTATAGATGGAGTTTGTATGAGTGGGGATAAAACAATGTACAGACACATTT 1568  
DB 481 GlnTrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPhe 500  
QY 1569 GGCAAAAGATAGAGTTTGGCCTTCAGAAATGTTTATTCACCTTAATTTCAAAGACTGAT 1628  
DB 501 GlyLysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAsp 520  
QY 1629 GTATCTGGAACCTGCTATTAAGTATCAGGATACACAGCTGGAAGCTGCAATCTG 1688  
DB 521 ValSerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeu 540  
QY 1689 CAGTGGGAAAAATTAATATGTAATATGATAGGTAAATTTTTTATCAAAATTCGAAGGCC 1748  
DB 541 GlnCysGlyLysLeuIleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAla 560  
QY 1749 ACTATTATTATGCCAACATAAGTGGACATCTCTGCATTCCTGTGGAATTTGCCAGTGAT 1808  
DB 561 ThrIleIleTyrAlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAsp 580  
QY 1809 CATGCACAGACCCAAAGATGTGGATAAAGATGGAACCTCTTGTGTTGTTCAATTAAGTGT 1868  
DB 581 HisAlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysVal 600  
QY 1869 TGCAGGAATCAAAAGATGTGTGAGTTCTTTCATCTTGGGTTATGATTGTACTACTGACAAA 1928  
DB 601 CysArgAsnGlnArgCysValSerSerSerTyrLeuGlyTyrAspCysThrThrAspLys 620  
QY 1929 TGCATATGATAGAGTGTATGCAATAACAAAGACCTGTCTGCTAGTGTCTTCATATTTA 1988  
DB 621 CysAsnAspArgGlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeu 640  
QY 1989 CCTCCAGATGCTCAGTTCATCAGATCTATGCGCTGCTGGGAGTATTGACAGTGGCAAT 2048  
DB 641 ProProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsn 660  
QY 2049 TTTTCCACCTGTAGCTATACCAGCCAGACTCCCTCGAAAGCGCTACATTGTAGAACATTATAC 2108  
DB 661 PheProProValAlaIleProAlaArgLeuProGluArgArgTyrIleGluAsnIleTyr 680  
QY 2109 CATTCCAAACCAATGAGATGGCCATTTTCTTATTCATTCTTCTTTTATTTTCTGT 2168  
DB 681 HisSerLysProMetArgTrpProPhePheLeuPheIleProPhePheIleIlePheCys 700  
QY 2169 GTACTGATCTTATATGTTGAAGTAAATTTCCNAAGGAAAATAAGACACTGACGAC 2228  
DB 701 ValLeuIleAlaIleMetValLysValAsnPheGlnArgLysLysTrpArgThrGluAsp 720

QY 2229 TATTCAAGCGATGACCACTGAAAGTGGAGTGAACCTAAAGG 2273  
 Db 721 TyrSerSerAspGluGlnProGluSerGluSerGluProLysGly 735

RESULT 2  
 AD02\_MACFA STANDARD; PRT; 735 AA.  
 ID AD02\_MACFA Q28472;  
 AC 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)  
 DE (Fertilin beta subunit) (PH-30) (PH30) (PH30-beta).  
 GN Name=ADAM2; Synonyms=FTNB;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=95260313; PubMed=7741716;  
 RA Perry A.C.F., Gichuhi P.M., Jones R., Hall L.;  
 RT "Cloning and analysis of monkey fertilin reveals novel alpha subunit  
 isoforms.";  
 RL Biochem. J. 307:843-850(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96341733; PubMed=8720115;  
 RA Ramarao C.S., Myles D.G., White J.M., Primakoff P.;  
 RT "Initial evaluation of fertilin as an immunoreceptive antigen and  
 molecular cloning of the cynomolgus monkey fertilin beta subunit.";  
 RL Mol. Reprod. Dev. 43:70-75(1996).  
 CC -1- FUNCTION: Sperm surface membrane protein that may be involved in  
 sperm-egg plasma membrane adhesion and fusion during  
 fertilization. Could have a direct role in sperm-zona binding or  
 migration of sperm from the uterus into the oviduct. Interactions  
 with egg membrane could be mediated via binding between its  
 disintegrin-like domain to one or more integrins receptors on the  
 egg. This is a non catalytic metalloprotease-like protein (By  
 similarity).  
 CC -1- SUBUNIT: Heterodimer with ADAM1/fertilin alpha.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Expressed specifically in testis.  
 CC -1- DOMAIN: A tripeptide motif (PDE) within disintegrin-like domain  
 could be involved in the binding to egg integrin receptor and thus  
 could mediate sperm/egg binding (By similarity).  
 CC -1- PTM: The signal and the metalloprotease domain are cleaved during  
 the epididymal maturation of the spermatozoa (By similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin domain.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; X77653; CAAS4733.1; -;  
 DR EMBL; U33959; BAB52699.1; -;  
 DR PIR; G02937; G02937.  
 DR HSP; P30403; IN4Y.  
 DR MEROPS; M12.950; -;  
 DR InterPro; IPR006586; ADAM cysteine.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR000742; EGF 2.  
 DR InterPro; IPR006203; EGF-like.  
 DR InterPro; IPR001590; Peptidase\_M12B.

InterPro; IPR002870; Peptidase\_M12B\_N.  
 Pfam; PF00200; Disintegrin; 1.  
 Pfam; PF01562; Pep\_M12B\_propep; 1.  
 Pfam; PF01421; Reprolysin; 1.  
 DR PRINTS; PR00289; DISINTEGRIN.  
 DR PRODOM; PD000664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR PROSITE; PS02115; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; 1.  
 DR PROSITE; PS00214; DISINTEGRIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; FALSE NEG.  
 DR PROSITE; PS01186; EGF\_2; FALSE NEG.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 KW Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
 FT SIGNAL 1  
 FT PROPEP 17 174 By similarity.  
 FT CHAIN 175 735 ADAM 2.  
 FT DOMAIN 17 686 Extracellular (Potential).  
 FT TRANSMEM 687 707 Potential.  
 FT DOMAIN 708 735 Cytoplasmic (Potential).  
 FT DOMAIN 175 382 Metalloprotease-like.  
 FT DOMAIN 383 475 Disintegrin-like.  
 FT DOMAIN 476 606 Cys-rich.  
 FT DOMAIN 612 645 EGF-like.  
 FT DISULFID 370 370 By similarity.  
 FT DISULFID 329 354 By similarity.  
 FT DISULFID 445 458 Potential.  
 FT DISULFID 616 627 By similarity.  
 FT DISULFID 621 633 By similarity.  
 FT DISULFID 635 644 By similarity.  
 FT CARBOHYD 76 76 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 122 122 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 220 220 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 353 353 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 459 459 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 566 566 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 723 723 S -> T (in Ref. 2).  
 SQ SEQUENCE 735 AA; 82357 MW; BE84BE115E1D0B53 CRC64;

Alignment Scores:  
 Pred. No.: 1.53e-270 Length: 735  
 Score: 3690.50 Matches: 672  
 Percent Similarity: 95.51% Conservative: 30  
 Best Local Similarity: 91.43% Mismatches: 32  
 Query Match: 76.95% Indels: 1  
 DB: 1 Gaps: 1

US-10-054-683-18 (1-2640) x AD02\_MACFA (1-735)  
 QY 72 ATGTGG---GTCTTGTCTGCTCAGCGGGCTCGGGGCTGGGATGGACAGTAATTTT 128  
 Db 1 MetcTrpArgValLeuPheLeuSerGlyLeuGlyLeuTrpMetAspSerAsnPhe 20  
 QY 129 GATAGTTTACCTGTGCAAAATTACAGTTCGGAGAAATACGGTCAATAATAAGGAAGGA 188  
 Db 21 AspSerLeuProValGlnIleThrValProGluLysIleArgSerIleLeuGluGlu 40  
 QY 189 ATTGAATCCAGGCATCCTACAAAATTGTAATTGAAGGGAACCATATATCTGTGAATTTA 248  
 Db 41 IleGluSerGlnValSerTrpLysIleValIleGluGlyLysProTrpThrAlaAsnLeu 60  
 QY 249 ATGCAAAAAAATCTTTTACCCTAATTTTACAGTTTACAGTTATAGTCGCACAGGAAT 308  
 Db 61 MetGlnLysAsnPheLeuSerHisAsnPheArgValTrpSerTrpAsnGlyThrGlyIle 80  
 QY 309 ATGAAACCACTTGACCAAGATTTCAGCAATTTCTGCACCTACCAGGGTATATTGAAGGT 368  
 Db 81 MetLysProLeuAspGlnAspPheGlnAsnPheCysHisTrpGlnGlyTrpIleGluGly 100  
 QY 369 TATCCAAAATCTGT 428  
 Db 101 TyrProLysSerValAlaMetValSerThrCysThrGlyLeuArgGlyLeuGlnPhe 120



RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 DR ENBL; BC064547; AAH64547.1; -  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR Pfam; PF00200; Disintegrin; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propag; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR PRINTS; PD00289; DISINTEGRIN.  
 DR PRODOM; PD00664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR PROSITE; PS00215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; 1.  
 DR PROSITE; PS00214; DISINTEGRIN\_2; 1.  
 SQ SEQUENCE 579 AA; 64795 MW; 1F7AF0CDA2A064F9 CRC64;

Alignment Scores:  
 Pred. No.: 2,97e-233 Length: 579  
 Score: 3065.50 Matches: 577  
 Percent Similarity: 78.50% Conservative: 0  
 Best Local Similarity: 78.50% Mismatches: 1  
 Query Match: 63.92% Indels: 157  
 DB: 2 Gaps: 3

US-10-054-683-18 (1-2640) x Q6P2G0 (1-579)

QY 72 ATGTGG---GTCCTTGTCTCTCAGCGGGCTCGCGGGCTCGGATGGACAGTAATTTT 128  
 DB 1 MetTrpArgValLeuPheLeuSerGlyLeuGlyGlyLeuArgMetAspSerAsnPhe 20  
 QY 129 GATAGTTTACTGTGCAAAATTTACAGTTCGGAGAAATACGGTCAATAATAAGGAAGGA 188  
 DB 21 AspSerLeuProValGlnIleThrValProGlnLysIleArgSerIleLeuLysGluGly 40  
 QY 189 ATTCAATCGCAGGATCTCTACAAATTTGTAATTAAGGAAACCATATCTACTGTGAATTTA 248  
 DB 41 IleGluSerGlnAlaSerTyrLysIleValIleGluLysProTyrThrValAsnLeu 60  
 QY 249 ATGCAAAAAAATTTTACCCCAATTTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTT 308  
 DB 61 MetGlnLysAsnPheLeuProHiaAsnPheArgValTyrSerTyrSerGlyThrGlyLe 80  
 QY 309 ATGAACACCTTGACCAAGATTTTTCAGAAATTTCTGCCACTACCAAGGTATATTGAAGGT 368  
 DB 81 MetLysProLeuAspGlnAspPheGlnAsnPheCysHisIleTyrGlnGlyTyrIleGluGly 100  
 QY 369 TATCCAAAATCTGGGTGATGGTATGACACATCTACTTGGACTCAGGGCGGTACTACAGTTT 428  
 DB 101 TyrProLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPhe 120  
 QY 429 GAAATGTTAGTTAGTAATAGAACCCCTCGAGTCTTCAGTTGGCTTTGAACATGTAATT 488

DB 121 GluAsnValSerTyrGlyIleGluProLeuGluSerSerValGlyPheGluHisValIle 140  
 QY 489 TACCAAGTAAACATAGAAAGCAGATGTTCTTATATATATATAGAAAGGATATTGAATCA 548  
 DB 141 TyrGlnValLysHisLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 160  
 QY 549 AGAGATCTCTCTCTTAAATTTACAAAGCCGAGAGCCACAGCAAGATTTTTCAAAGATATA 608  
 DB 161 ArgAspLeuSerPheLysLeuGlnSerValGlu----- 171  
 QY 609 GAAATGCATGTTATAGTTGAAAAACAANTTGTATATATCATATGGGCTCTGATACAACTGTT 668  
 DB 171 ----- 171  
 QY 669 GTCGCTCAAAAAGTTTTCAGTTGATTGATTGACGAATGCTATTGTTGTTTCATTAAAT 728  
 DB 171 ----- 171  
 QY 729 ATTACAATTTATCTGCTTTCATTTGGAGCTTTGGATAGATGAATAAATTCGAACCACT 788  
 DB 171 ----- 171  
 QY 789 GGAGAAGCTAATGAGTTATTACACATTTTAAAGATGGAAAACATCTTATCTGTTTAA 848  
 DB 171 ----- 171  
 QY 849 CGTCCTCATGATGTGGCATTGTTTACTGTTTACAGAGAAAAGTCAAAATTTATGTTGTCGA 908  
 DB 171 ----- 171  
 QY 909 ACCTTTCAAGGGAAGATGTGTGATGCAAACTATGAGGAGGTGTTGTTCTGCNCCCAGCA 968  
 DB 172 -----HisProArg 174  
 QY 969 ACCATAAGTCTGGAATCACTTCAGTTATTTTAGCTCAATTTATTGAGCCTTAGTATGGGG 1028  
 DB 175 ThrIleSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGly 194  
 QY 1029 ATCATTATGATGACATTAACAAATGCCAGTGTCTCAGGAGCTGTCTGCATTATGAATCCA 1088  
 DB 195 IleThrTyrAspAspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnPro 214  
 QY 1089 GAAGCAATTCATTTTCACTGAGTGAAGATCTTTAGTAACTGACGCTTCCAGAGCTTTGCA 1148  
 DB 215 GluAlaIleHisPheSerGlyValLysIlePheSerAsnCysSerPheGluAspPheAla 234  
 QY 1149 CATTATTTTCAAGCAGAGTCCCAAGTGTCTTCACAATCAGCTCGCTTAGATCTCTTTT 1208  
 DB 235 HisPheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhe 254  
 QY 1209 TTCAAAACAGCAGCAGTGTGTGTAATGCAAGCTCGAAGCAGAGAGAGAGTGTGACTGT 1268  
 DB 255 PheLysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCys 274  
 QY 1269 GGACTGAAACAGATTTGTCCTTATTGGAAGAACATGCTGTCATATTCACATGATGAGA 1328  
 DB 275 GlyThrGluGlnAspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArg 294  
 QY 1329 TTTTAAAGCCGTTTCAAACTGTGCTGAAGGACCATGCTGCGAAAACCTGCTTATTATGTCA 1388  
 DB 295 PheLysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSer 314  
 QY 1389 AAAGAAGAATGTGTAGGCTTCTCTTTGAAGAAATGGACCTCCCTCGAATATTCGAATGGA 1448  
 DB 315 LysGluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGly 334  
 QY 1449 TCATCTGCATCATGCCCGAAGAACCTATGTTTCAGACTGGGATCCGCTGTCGACTGAAT 1508  
 DB 335 SerSerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsn 354  
 QY 1509 CAATGGATCTGTATAGATGAGTGTGTATGAGTGGGATGAACAATGTACAGACATTT 1568  
 DB 355 GlnTrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPhe 374

QY 1569 GGCAGAGAGTAGAGTTGGCCCTTCAGAAATGTTATTTCTCACCTTAATTTCAAGACTGAT 1628  
D 375 GlyLysGluValGluPheGlyProSerGluCysTyrSerHisLeuAenSerLysThrAsp 394  
QY 1629 GTATCTGGAACCTGGTATAGTATTAGGATACACACAGCTGTAAGCTGCAATCTG 1688  
D 395 ValSerGlyAsnGlyHisSerAspSerGlyTyrThrGlnCysGluAlaAsp----- 412  
QY 1689 CAGTGGGAAATTAATATGTAATATGAGTAAATTTTATTACAAATTTCAAGAGCC 1748  
D 412 ----- 412  
QY 1749 ACTATTATTATGCCAACATAAGTGGACATCTCTGCATTCCTGTGGAATTTGCCAGTGAT 1808  
D 413 -----GlyHisLeuCysAlaValGluPheAlaSerAsp 424  
QY 1809 CATGCAGACAGCCAAAAGATGTGGATAAAGATGGAACCTCTTGTGGTTCAAATAAGTT 1868  
D 425 HisAlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysVal 444  
QY 1869 TGCAGGAATCAAGATGTGTGAGTCTTCTCATCTTGGTTATGATGTACTACGACAAA 1928  
D 445 CysArgAsnGlnArgCysValSerSerTyrLeuGlyTyrAspCysThrThrAspLys 464  
QY 1929 TGCATATGATAGAGTGTATGCAATAACAAAGACACTGTCACTGTAGTGTCTCATATTTA 1988  
D 465 CysAsnAspArgGlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeu 484  
QY 1989 CTTCCAGATTGCTCAGTTCAATCAGATCTATGCGCTGGTGGAGTATTGACAGTGGCAAT 2048  
D 485 ProProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsn 504  
QY 2049 TTTCCACCTGACTATACAGCCAGACTCCCTGAAAGGCGTACATTGAGAACATTAC 2108  
D 505 PheProProValAlaIleProAlaLysGluProGluArgTyrIleGluAsnIleTyr 524  
QY 2109 CATTCACAAACCAATGAGATGGCCATTTTCTTATTCATCTCTTCTTATTTATTTCTGT 2168  
D 525 HisSerLysProMetArgTrpProPhePheLeuPheIleProPhePheIlePheCys 544  
QY 2169 GTACTGATCTCTATATGTTGAAGTTAATTTCAAGGAAAAAATGAGAACTGAGGAC 2228  
D 545 ValIleAlaIleMetValLysValAsnPheGlnArgLysLysTrpArgThrGluAsp 564  
QY 2229 TATTCAAGCATGAGCAACTGAAAGTGAGAGTGAAGCTAAAGGG 2273  
D 565 TyrSerSerAspGluGlnProGluSerGluSerGluProLysGly 579  
RESULT 4  
AAH64547  
ID AAH64547 PRELIMINARY; PRT; 579 AA.  
AC AAH64547;  
DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
DE ADAM2 protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC064547; AAH64547.1; -. 1F7AP0CDA2A064F9 CRC64;  
SQ SEQUENCE 579 AA; 64795 MW; 1F7AP0CDA2A064F9 CRC64;

## Alignment Scores:

Pred. No.: 2,97e-223 Length: 579  
Score: 3055.50 Matches: 577  
Percent Similarity: 78.50% Conservative: 0  
Best Local Similarity: 78.50% Mismatches: 1  
Query Match: 63.92% Indels: 157  
DB: 2 Gaps: 3

US-10-054-683-18 (1-2640) x AAH64547 (1-579)

QY 72 ATGTGG---GTCTGTTTCTGCTCAGCGGCTCGCGGCTCGGATGGACAGTAATTTT 128  
D 1 MetTrpArgValLeuPheLeuLeuSerGlyLeuGlyLeuArgMetAspSerAsnPhe 20  
QY 129 GATAGTTTACCTGTGCAATTTACAGTCCGAGAAAATACGTCATATAAAGGAGGA 188  
D 21 AspSerLeuProValGlnIleThrValProGluLysIleArgSerIleLysGluGly 40  
QY 189 ATTGAATCGCAGCATCTCACAATAATTGAATTTGAAGGGGAAACATATCTCTGTAATTTA 248  
D 41 IleGluSerGlnAlaSerTyrLysIleValIleGluGlyLysProTyrThrValAsnLeu 60  
QY 249 ATGCAAAAAAATTTTACCCATAATTTTAGAGTTTACAGTTTATAGTGCACAGGAAT 308  
D 61 MetGlnLysAsnPheLeuProHisAsnPheArgValTyrSerTyrSerGlyThrGlyLe 80  
QY 309 ATGAAACCACTTGCACCAAGATTTTCAGATTTTCGCACCTACCAAGGCTATATGAGGT 368  
D 81 MetLysProLeuAspGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGly 100  
QY 369 TATCCAAAATCTGTGTTGATGGTTAGCACATGTACTGGACTCAGGGCGGTACTACAGTTT 428  
D 101 TyrProLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPhe 120  
QY 429 GAAAAATGTTAGTTATGGAATAGAACCCCTGGAGTCTTTCAGTTCGCTTTGAACATGTAAT 488  
D 121 GluAsnValSerTyrGlyIleGluProLeuGluSerSerValGlyPheGluHisValIle 140  
QY 489 TACCAAGTAAACATAGAAGCAGATGTTTCCTTATATATAGAGGATATATGATCA 548  
D 141 TyrGlnValLysHisLysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSer 160  
QY 549 AGAGATCTGTCTTTAAATTTACAAAGCGCAGAGCCACAGCAAGATTTTGCAGAGTATATA 608  
D 161 ArgAspLysSerPheLysLeuGlnSerValGlu----- 171  
QY 609 GAAATGCATGTTTAGTTGAAAAACAAATGTTATATCATATGGGCTCTGATACAACATGTT 668  
D 171 ----- 171  
QY 669 GTCGCTCAAAAAGCTTTTCCAGTTGATTTGATGACGAAATGCTATTTTGTTCATTTAAT 728  
D 171 ----- 171



QY 729 ATTACAAATTATTCTGCTTCAATGGAGCTTTGGTAGATGAAATTAATAATGCAACCACT 788  
Db 171 ----- 171  
QY 789 GGAGAAGCTAATGAGTTATTACACACATTTTAAAGATGGAAACATCTTATCTTGTGTTTA 848  
Db 171 ----- 171  
QY 849 CGTCCTCATGATGGCAATTTTACTGTTTACAGAGAAAGTCAAAATATGTTGGTGCA 908  
Db 171 ----- 171  
QY 909 ACCTTTCAAGGGAAGATGTGTGATGCAAACTATGCAGGAGGTGTGTTCTGCAACCCGAGA 968  
Db 172 -----HisProArg 174  
QY 969 ACCATAAGTCTGGAATCAGTTGAGTTATTATTAGCTCAATATTAGCCCTTAGTATGGG 1028  
Db 175 ThrIleSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerMetGly 194  
QY 1029 ATCACTTATGATGACATTACAAATGCCAGTGTCTCAGAGCTGTCTGCATTATGATCCA 1088  
Db 195 IleThrTyrAspAspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnPro 214  
QY 1089 GAAGCAATTCATTTACAGTGTGTGAAGATCTTTAGTAACTGCACTTCGAGAGCTTTGCA 1148  
Db 215 GluAlaIleHisPheSerGlyValLysIlePheSerAsnCysSerPheGluAspPheAla 234  
QY 1149 CATTTTATTTCAAGCAGAAAGTCCAGTGTCTTCACAATCAGCTCGCTTAGATCCCTTT 1208  
Db 235 HisPheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhe 254  
QY 1209 TTCAACAGCAGCAGTGTGTGTAATGCAAGCTGCAAGCAGAGAGGAGTGTGACTGT 1268  
Db 255 PheLysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluCysAspCys 274  
QY 1269 GGGAGTCAACAGGATTTGGCCCTTATTGGAGAAACATGCTGTGATATTGGCACATGTAGA 1328  
Db 275 GlyThrGluGlnAspCysAlaLeuIleGlyLysCysAspIleAlaThrCysArg 294  
QY 1329 TTTAAAGCCGGTCAAACTGTGCTGAAGGACCATGCTGCGGAAACCTGCTATTATGTCA 1388  
Db 295 PheLysAlaGlySerAsnCysAlaGlyProCysCysGluAsnCysLeuPheMetSer 314  
QY 1389 AAAGAAAGAAATGCTAGCCCTTCTTGAAGAAATGCGACCTCCCTGATATTTGCAATGGA 1448  
Db 315 LysGluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGly 334  
QY 1449 TCATCTGCATCATGCCAGAAACCATCATGTTTCAAGTGGGCACTCCGCTGTGGACTGAAT 1508  
Db 335 SerSerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsn 354  
QY 1509 CAATGGATCTGTATAGATGGAGTTTGTATAGTGGGGATAAACAAATGTACAGACACATTT 1568  
Db 355 GlnTrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPhe 374  
QY 1569 GGCAAGAGATAGATTTGGCCCTTCAAGATGTTATTTCTACCTTAATTTCAAGACTGAT 1628  
Db 375 GlyLysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAsp 394  
QY 1629 GTATCTGAAACTGTGGTATAGTATAGTGGGATTAACAAATGTACAGACATCTG 1688  
Db 395 ValSerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAsp----- 412  
QY 1689 CAGTCGGGAAATTAATATGTAATATGATAGTAAATTTTATTACAAATTTCCAGAGGCC 1748  
Db 412 ----- 412  
QY 1749 ACTATTATTATGCCAACATAGTGGACATCTCTGATTTGCTGTTGGAATTTGCCAGTGTAT 1808  
Db 413 -----GlyHisLeuCysIleAlaValGluPheAlaSerAsp 424

QY 1809 CATGCAGACAGCAAAAGATGTGGATAAAGATGAACCTCTTGTGGTTCAAAATAGGTT 1868  
Db 425 HisAlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysVal 444  
QY 1869 TCAGGAATCAAGAGTGTGAGTCTTTCATACTTGGGTTATGATTGTACTACTACAAA 1928  
Db 445 CysArgAsnGlnArgCysValSerSerSerTyrLeuGlyTyrAspCysThrThrAspLys 464  
QY 1929 TGCAATGATAGAGGTGTATGCAATAACAAAAAGCACTGTCACTGTAGTCTTCAATTTA 1988  
Db 465 CysAsnAspArgGlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeu 484  
QY 1989 CTCCAGATTGCTCAGTTCAATCAGATCATATGCGCTGTGGAGTATTGACAGTGGCAAT 2048  
Db 485 ProProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsn 504  
QY 2049 TTTCCACCTGTAGCTATACAGCCAGCTCCCTGAAAGCGCTACATTGAGAACATTAC 2108  
Db 505 PheProProValAlaIleProAlaArgLeuProGluArgTyrIleGluAsnIleTyr 524  
QY 2109 CATTTCCAAACCAATGAGATGGCCATTTTCTTATTTCATTCTCTTATTATTCTCTGT 2168  
Db 525 HisSerLysProMetArgTrpPhePheLeuPheIleProPhePheIleIlePheCys 544  
QY 2169 GTACTGATTGCTATATGTTGAAGTTAATTTCCAAAGGAAAAATGGAGAACTGAGGAC 2228  
Db 545 ValLeuIleAlaIleMetValLysValAsnPheGlnArgLysLysTyrArgThrGluAsp 564  
QY 2229 TATTCAGCGATCAGCAACCTGAAAGTGAAGTGAACCTTAAAGGG 2273  
Db 565 TyrSerSerAspGluGlnProGluSerGluSerGluProLysGly 579  
RESULT 5  
Q866A8 PRELIMINARY; PRT; 735 AA.  
ID Q866A8  
AC Q866A8; 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Pertilin beta.  
GN Names=FTNB;  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Butharia; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=22872989; PubMed=14510675;  
RA Day A.E., Quilter C.R., Sargent C.A., Mileham A.J.;  
RT "Chromosomal mapping, sequence and transcription analysis of the  
porcine fertilin beta gene (ADAM2).";  
RL Anim. Genet. 34:375-378(2003).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Day A.;  
RL Thesis (2004), Department of Applied Sciences, Anglia Polytechnic  
University, Cambridge, United Kingdom.  
RL EMBL; AJ309003; CAC84225.2; --  
DR HSP; P18619; IPVL.  
DR MEROPS; M12.950; --  
DR GO; GO:0004222; F:metallopeptidase activity; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR006586; ADAM\_cysteine.  
DR InterPro; IPR001762; Disintegrin.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR008994; Nucleic\_acid\_OR.  
DR InterPro; IPR001590; Peptidase\_M12B.  
DR InterPro; IPR002870; Peptidase\_M12B\_N.  
DR Pfam; PF00200; Disintegrin; 1.





modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

CC EMBL; AF086808; AAC62753.1; --  
 CC DR HSSP; P18619; 1FVL.  
 CC DR MEROPS; M12.950; --  
 CC DR InterPro; IPR006586; ADAM cysteine.  
 CC DR InterPro; IPR001762; Disintegrin.  
 CC DR InterPro; IPR000742; EGF\_2.  
 CC DR InterPro; IPR006209; EGF\_like.  
 CC DR InterPro; IPR006210; IEGF.  
 CC DR InterPro; IPR008994; Nucleic acid OB.  
 CC DR InterPro; IPR001590; Peptidase M12B.  
 CC DR InterPro; IPR002870; Peptidase M12B\_N.  
 CC DR Pfam; PF00200; Disintegrin; 1.  
 CC DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 CC DR Pfam; PF01421; Reprolysin; 1.  
 CC DR PRINTS; PR00289; DISINTEGRIN.  
 CC DR ProDom; PD000664; Disintegrin; 1.  
 CC DR SMART; SM00608; ACR; 1.  
 CC DR SMART; SM00050; DISIN; 1.  
 CC DR SMART; SM00181; EGF; 1.  
 CC DR PROSITE; PS00215; ADAM\_MBP; 1.  
 CC DR PROSITE; PS00427; DISINTEGRIN\_1; 1.  
 CC DR PROSITE; PS00214; DISINTEGRIN\_2; 1.  
 CC DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 CC DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 CC DR PROSITE; PS00026; EGF\_3; 1.  
 CC DR Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
 CC FT SIGNAL 1 16 Potential.  
 CC FT PROPEP 17 174 By similarity.  
 CC FT CHAIN 175 745 ADAM 2.  
 CC FT DOMAIN 17 696 Extracellular (Potential).  
 CC FT TRANSMEM 697 717 Potential.  
 CC FT DOMAIN 718 745 Cytoplasmic (Potential).  
 CC FT DOMAIN 175 382 Metalloprotease-like.  
 CC FT DOMAIN 383 474 Disintegrin-like.  
 CC FT DOMAIN 475 605 Cys-rich.  
 CC FT DOMAIN 611 644 EGF-like.  
 CC FT DOMAIN 148 151 Poly-Ser.  
 CC FT DISULFID 287 370 By similarity.  
 CC FT DISULFID 329 354 By similarity.  
 CC FT DISULFID 444 457 Potential.  
 CC FT DISULFID 615 626 By similarity.  
 CC FT DISULFID 620 632 By similarity.  
 CC FT DISULFID 634 643 By similarity.  
 CC FT CARBOHYD 122 122 N-linked (GlcNAc...) (Potential).  
 CC FT CARBOHYD 220 220 N-linked (GlcNAc...) (Potential).  
 CC FT CARBOHYD 353 353 N-linked (GlcNAc...) (Potential).  
 CC FT CARBOHYD 458 458 N-linked (GlcNAc...) (Potential).  
 CC FT CARBOHYD 558 558 N-linked (GlcNAc...) (Potential).  
 CC SQ SEQUENCE 745 AA; 83150 MW; B5D8DC0168999B00 CRC64;

Alignment Scores:  
 Pred. No.: 2,51e-197 Length: 745  
 Score: 2723.00 Matches: 481  
 Percent Similarity: 78.71% Conservative: 103  
 Best Local Similarity: 64.82% Mismatches: 146  
 Query Match: 56.78% Indels: 12  
 DB: 1 Gaps: 2

US-10-054-683-18 (1-2640) x AD02\_BOVIN (1-745)

QY 78 GTCTGTTTCTGCTCAGCGGGCTCGCGGGCTCGCGGAGTACAGTAAATTTTGTAGTTTA 137  
 Db 4 leuLeuPheLeuLeuLeuGlyLeuThrGlyLeuThrAspAspAsnSerGluArgLeu 23

QY 138 CCTGTGCAATTTACGTTCCGGAGAAATACGTTCAATATAAAGGAAGGAATTTGAATCG 197  
 Db 24 ArgValGlnPheThrValProGluLysIleArgSerThrSerGlyGlyValGluThr 43

QY 198 CAGGCATCTCACAAATTTGTAATTTGAGGGGAACCATATATCTGTGAATTTTAATGCAAAA 257

1938 AGAGGTGATGCAATAACAAAAGCACTGTCACGTAGTGTCTCATATTTTACCTCCAGAT 1997  
 Db 623 GlnGlyValCysAsnAsnLysArgAsnCysHisCysLysProThrThrLeuProProAsn 642

1998 TGCTCAGTTCAATCAGATCATATGCGCTGGTGGGAGTATTGACAGTGGCAATTTTCCACCT 2057  
 Db 643 CysGluValSerValGluSerTrpProGlyGlySerValAspSerGlyAsnPheProPro 662

2058 GTAGCTATACA---GCCAGACTCCCTGAAGGCGGTACATTTGAGACATTTACCATTC 2114  
 Db 663 AlaAlaAspProLeuProGlyValProAspArgArgTyrIleGluAsnIleTyrHisSer 682

2115 AAACCAATGAGATGCGCATTTTCTTATTCATTCCTTTCTTTATTTATTTCTGTGTACTG 2174  
 Db 683 ArgProMetLysTrpProPhePheLeuLeuIleProPhePheIleLeuCysValLeu 702

2175 ATTGCTATAATGGTGAAGTTAAATTTCCAAAGGAAAAATGAGAACTGAGGACTATTCA 2234  
 Db 703 IleAlaThrLeuValLysIleTyrPheGlnArgLysLysGlyArgThrGluAspAspThr 722

2235 AGCGATGAGCACTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2270  
 Db 723 SerAspGlnGlnLeuGluSerGluSerGluSerGluThrLys 734

RESULT 6  
 AD02\_BOVIN  
 ID \_AD02\_BOVIN STANDARD; PRT; 745 AA.  
 AC 077780;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)  
 DE (Fertilin beta subunit) (PH-30) (PH30) (PH30-beta).  
 GN Name=ADAM2; Synonyms=FTNB;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=97304362; PubMed=9160725;  
 RA Waters S.I., White J.M.;  
 RT "Biochemical and molecular characterization of bovine fertilin alpha  
 and beta (ADAM 1 and ADAM 2): a candidate sperm-egg binding/fusion  
 complex.";  
 RL Biol. Reprod. 56:1245-1254(1997).  
 CC -1- FUNCTION: Sperm surface membrane protein that may be involved in  
 sperm-egg plasma membrane adhesion and fusion during  
 fertilization. Could have a direct role in sperm-zona binding or  
 migration of sperm from the uterus into the oviduct. Interactions  
 with egg membrane could be mediated via binding between its  
 disintegrin-like domain to one or more integrins receptors on the  
 egg. This is a non catalytic metalloprotease-like protein (By  
 similarity).  
 CC -1- SUBUNIT: Heterodimer with ADAM1/fertilin alpha.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Expressed specifically in testis.  
 CC -1- DOMAIN: A tripeptide motif (TDE) within disintegrin-like domain  
 could be involved in the binding to egg integrin receptor and thus  
 could mediate sperm/egg binding (By similarity).  
 CC -1- PTM: The signal and the metalloprotease domain are cleaved during  
 the epididymal maturation of the spermatozoa (By similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin domain.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way

44	HisValSerTyrIleLeuIleGluGlyThrTyrThrValAsnLeuMetGlnLys	63
258	AACTTTTACCCTAATTTTACAGTTTACAGTTTATAGTGCACACAGAAATTATGAACCA	317
64	AlaPheLeuProHisAsnPheArgValTyrGlyTyrSerGlyThrGlySerMetLysPro	83
318	CTTGACCAAGATTTTACAGATTTCTGCCACTACCAAGGGTATATTGAAGTTTATCAAAA	377
84	LeuGluHisGluPheGlnAsnPheCysTyrTyrGlnGlyTyrIleGluGlyTyrProAsn	103
378	TCGTGTGTGGTTAGTACCATGATCTGAGACTCAGGGCGCTACTACAGTTTGAATGTT	437
104	SerMetAlaIleIleSerThrCysThrGlyLeuArgGlyLeuLeuGlnPheGluAsnVal	123
438	AGTTATGGAATGAACCCCTGGAGTCTTCAGTTGGCTTCAACATGATTAATTTACCAAGTA	497
124	SerTyrGlyIleGluProLeuGluProSerIleGlyPheGluHisMetValTyrGlnIle	143
498	AAACATGAAGAAAGCAGATGTTTCCTTATATAATCAGAAGGATATTGAATCAAGAGATCTG	557
144	LysProArgAspSerSerSerValTyrThrGluArgGluIleGluLeuArgGluLys	163
558	TCCTTTAAATTAACAAGCGCAGACGACAGAGATTTTGCAAAGTATATAGAAATGTCAT	617
164	ProTyrIysIleGlnAsnValGluProLeuProAspPheSerGlnTyrIleGluMetHis	183
618	GTTATAGTTCAAAAACAATTTGATATCATATCGGTCTGATCACTGCTGTTGTCGCTCAA	677
184	IleValValGluLysAspLeuTyrAsnHisMetGlyAlaAspThrThrValValIleGln	203
678	AAAGTTTTCAGATGATTGGATTCACCAATGCTATTTTGTTCATTATATTAACAATT	737
204	LysIlePheGlnLeuThrGlyLeuThrAsnAlaIlePheThrSerLeuAsnIleThrVal	223
738	ATTCTGTCTCATTTGGAGCTTTGGATAGATGAATAAAATTCGACCACTCGGAGAGCT	797
224	IleLeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleProValThrGlyAspAla	243
798	AATGAGTATTACACACATTTTTAAGATGGAAACATCTTATCTGTGTTTACGTCCTCAT	857
244	AsnGluLeuLeuHisArgPheValLysTrpLysArgSerTyrLeuValLeuArgProHis	263
858	GATGTGCATTTTACTTGTTTACAGAGAAAAGTCAAAATTATGTTGTGCAACCTTTCAA	917
264	AspMetAlaPheLeuLeuValTyrArgGluLysSerAsnTyrIleGlyAlaThrPheGln	283
918	GGGAAGATGTGTGATGCAAACTATGCAGGAGGTGTGTTCTGCAACCCAGAACCATAAGT	977
284	GlyArgMetCysAspLysHisTyrGlyGlyValAlaLeuHisSerSerThrIleSer	303
978	CTGGAATCACTTGAGTATTATTAGCTCAATTATTAGCGCTTAGTAGGGGATCACTTAT	1037
304	LeuGluSerLeuAlaValIleIleIleGlnLeuLeuSerLeuSerMetGlyIleProTyr	323
1038	GATGACATTAACAATGCCAGTGTCTCAGGAGCTGTCTGCACTTATGAATCAGAAGCAATT	1097
324	AspAspIleAsnLysCysHisCysProGlyAspValCysIleMetAsnProAlaAlaVal	343
1098	CATTTCAGTGGTGAAGATCTTTTAGTAACTGCAGCTTCGAAGACTTTGACATTTTATT	1157
344	HisSerSerGlyValLysLeuPheSerAsnCysSerValGluAspPheLeuArgPheIle	363
1158	TCAAGCAGAGATCCCGATGCTTCACATCAGCGCTCGCTTAGTATCTTTTTTCACACAG	1217
364	SerLysProLysSerGlnCysLeuGlnAsnGlnProArgLeuAspProThrTyrLys---	382
1218	CAACGAGTGTGTGTAATGCAAGCTCGAAGCAGAGAGAGTGTGCTACTGTGGAGCTGAA	1277
383	SerAlaValCysGlyIleAsnGlyLysValGluGluGlnCysAspCysGlyAsnLys	402
1278	CAGGATTGTGCCCTTTATGGGAACAACATGCTGTGTGATATTGCCATGTAGATTAAAGCC	1337





CC -|- DOMAIN: A tripeptide motif (NQE) within disintegrin-like domain  
 CC could be involved in the binding to egg integrin receptor and thus  
 CC could mediate sperm/egg binding (By similarity).  
 CC -|- PTM: The prodomain and the metalloprotease domain are cleaved  
 CC during the epididymal maturation of the spermatozoa (By  
 CC similarity).  
 CC -|- SIMILARITY: Belongs to peptidase family M12B.  
 CC -|- SIMILARITY: Contains 1 disintegrin domain.  
 CC -|- SIMILARITY: Contains 1 EGF-like domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

DR EMBL; X99794; CAA68127.1; ALT\_INIT.  
 DR HSSP; P18619; LFVL.  
 DR MEROPS; M12.950; -.  
 DR RGD; 69299; Adam2.  
 DR InterPro; IPR006586; ADAM cysteine.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR000742; EGF 2.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR008994; Nucleic acid OB.  
 DR InterPro; IPR001590; Peptidase M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR Pfam; PF00200; Disintegrin; 1.  
 DR Pfam; PF01582; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR ProDom; PD000664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS50215; ADAM\_MEROP; 1.  
 DR PROSITE; PS00247; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PS50026; EGF\_3; 1.  
 KW Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
 FT SIGNAL 1 18 Potential.  
 FT PROPEP 19 180 By similarity.  
 FT CHAIN 181 737 ADAM 2.  
 FT DOMAIN 19 688 Extracellular (Potential).  
 FT TRANSMEM 689 709 Potential.  
 FT DOMAIN 710 737 Cytoplasmic (Potential).  
 FT DOMAIN 184 381 Metalloprotease-like.  
 FT DOMAIN 389 481 Disintegrin-like.  
 FT DOMAIN 482 616 Cys-rich.  
 FT DOMAIN 617 650 EGF-like.  
 FT DISULFID 293 376 By similarity.  
 FT DISULFID 335 360 By similarity.  
 FT DISULFID 441 457 Potential.  
 FT DISULFID 621 632 By similarity.  
 FT DISULFID 626 638 By similarity.  
 FT DISULFID 640 649 By similarity.  
 FT CARBOHYD 128 128 N-linked (GlcNAc...) (potential).  
 FT CARBOHYD 226 226 N-linked (GlcNAc...) (potential).  
 FT CARBOHYD 359 359 N-linked (GlcNAc...) (potential).  
 FT CARBOHYD 464 464 N-linked (GlcNAc...) (potential).  
 FT CARBOHYD 491 491 N-linked (GlcNAc...) (potential).  
 FT CARBOHYD 571 571 N-linked (GlcNAc...) (potential).  
 SQ SEQUENCE 737 AA; 82310 MW; 647E92A90D7CC1D1 CRC64;

## Alignment Scores:

Pred. No.: 6, 78-179 Length: 737  
 Score: 2479.50 Matches: 442  
 Percent Similarity: 75.30% Conservative: 116  
 Best Local Similarity: 59.65% Mismatches: 170

Query Match:	51.70%	Indels:	13
DB:	1	Gaps:	6
US-10-054-683-18 (1-2640) x AD02_RAT (1-737)			
QY	72	ATGTGGGTCTTGTCTCTG	-----CTCAGCGGGCTCGCGGGCTG----- 110
DB	1	MetTrpLeuLeuLeuLeuLeuSerGlyLeuSerArgLeuGlyGlyLeuSerGluPro	20
QY	111	CGGATGGACAGTAATTTTCATAGTTTACCTGTGCAATTTACAGTTCCGGAGAAATACGG	170
DB	21	GlnThrGluGlyThrArgGluLeuHisValGlnValThrValProGluLeuArg	40
QY	171	TCAAATAATAAGGAAGGAATTCGAATCGCAGGCATCTCTACAAAATGTAAATCAAGGAAA	230
DB	41	SerIleThrSerGluGlyGlyThrGlnThrGlnValThrSerLeuLeuGlyGlyLeu	60
QY	231	CCATATCTGTGAATTTAATGCAAAAAAATTTTACCCCAATAATTTTAGAGTTTACAGT	290
DB	61	ThrTyrIleLeuAsnLeuMetGlnLysAlaPheLeuProAsnPheArgValTyrSer	80
QY	291	TATAGTGGCACAGGAATTTATGAACCACTTGACCAAGATTTTCAGAATTTCTGCCACTAC	350
DB	81	TyrAspSerThrGlyIleMetArgProLeuGluGlnLysPheGlnAsnIleCysTyrPhe	100
QY	351	CAAGGGTATATTGAAGTTTATCCAAATCTGTGGTGATGGTTAGCACATGTACTGGACTC	410
DB	101	GlnGlyTyrIleGluGlyTyrProAsnSerMetValIleValSerThrCysThrGlyLeu	120
QY	411	AGGGCGTACTACAGTTTCAAAATGTTATGGAATAGAACCCCTGGAGTCTTCAGTT	470
DB	121	ArgGlyValLeuGlnPheGlyAsnValSerTyrGlyIleGluProLeuGluSerSer	140
QY	471	GGCTTTGAACATGTAATTTACCAAGTAAACATAGAAAGCAGATGTTTCCCTATATAAT	530
DB	141	GlyPheGluHisValIleTyrGlnValGluProLysLysGlyAspThrLeuLeuTyrAla	160
QY	531	GAGAAGCATATTGAATCAAGAGATCTGCTCTTAAATTAATAAGCGCAGCCACACAA	590
DB	161	GluLysAspMetAspLeuArgProGlnTyrLysIleArgSerIleLysProGlnArg	180
QY	591	GATTTTGAAGTATATAGAAATGATGATTTAGTTGAAACAAATGTTATATCATATG	650
DB	181	ThrValSerHisTyrLeuGluIleHisIleValValGluLysGlnMetPheGluHisIle	200
QY	651	GGTCTGTACACTGTTGTCGCTCAAAAGTTTCCAGTTGATTTGAGTTGAGCAATGCT	710
DB	201	GlyAlaAspThrAlaValThrGlnLysIlePheGlnLeuIleGlyLeuThrAsnAla	220
QY	711	ATTTTGTCTTCAATTAATTAATTAATTTCTCTCTTCTTCAATTTGGATAGTAA	770
DB	221	IlePheAlaProPheAsnLeuThrValIleLeuSerSerLeuGluPheTrpMetAspGlu	240
QY	771	ATAAAATTCGAACCACTGAGAGAGCTATAGATTATACACATTTTAAAGTGGAA	830
DB	241	AsnLysIleSerThrThrGlyAspAlaAsnLeuLeuTyrArgPheLeuTyrTrpLys	260
QY	831	ACATCTTCTGTTTCTGCTCTCATGATGTCGCTCAATTTTACTGTTTACAGAGAAAG	890
DB	261	GlnSerTyrLeuValLeuArgProHisAspMetAlaPheLeuLeuValTyrArgAspThr	280
QY	891	TCAAATATGTTGGTCAACCTTTCAAGGGAAGATGTGATGCAACATATGAGGAGGT	950
DB	281	ThrAspTyrValGlyAlaThrTyrGlnGlyLysMetCysAspLysAsnTyrAlaGlyGly	300
QY	951	GTTGTTCTGACCCCAAGAACCATAGTCGATCATCTTCAGTTTATTTAGTCAATTA	1010
DB	301	ValAlaLeuHisProLysAlaValThrLeuGluSerLeuAlaIleLeuValGlnLeu	320
QY	1011	TTGAGCTTATGATGGGATCACTTATGATGACATTAACAAATGCCAGTCTCTCAGGAGCT	1070
DB	321	LeuSerLeuSerMetGlyValAlaTyrAspAspValAsnThrCysGlnCysGlyValPro	340

QY 1071 GTCTGATTATGATTCAGAGCAATTCATTTCAGTGGTGTGAAGATCTTTAGTAAGTC 1130  
Db 341 ILeCysValMetAenProGluAlaLeuHisSerGlyValArgSerPheSerAsnCys 360  
QY 1131 AGCTTCGAAGACTTTGACACATTTTATTTCAAAGCAGAAAGTCCCACTGCTCTTCACAAATCAG 1190  
Db 361 SerMetGluAspPheSerLysPheIleValSerGlnSerSerHisCysLeuGlnAsnGln 380  
QY 1191 CTTGCTTAGATCTTTTCAACACAGCAAGCAGTGTGTGTAATGCAAGAGCTGGAAGCA 1250  
Db 381 ProHisLeuGlnProSerTyrLys---MetAlaValCysGlyAenGlyGluLeuGluGlu 399  
QY 1251 GGAGAGAGTGTGACTGCGGACTGAACAGGATGTGCCCTTATTGAGAGAAACATGCTGT 1310  
Db 400 GlyGluValCysAspCysGly---GlnGluGlyCysAspAspLysProProCysCys 418  
QY 1311 GATATTGCCACATGATAGATTAAAGCCGGTTCAAACCTGTGCTGAAGGACCATGTCGGAA 1370  
Db 419 AsnProThrThrCysGlnLeuSerGluGlySerThrCysSerThrGlySerCysAsp 438  
QY 1371 ---AAGTGTCTATTATGTCAAAGAAAGATGTGTAGGCTTCTTTTGAAGATGCGAC 1427  
Db 439 AlaSerCysAsnLeuLysAlaLysGlyGluLeuCysArgProAlaAsnGlnGluCysAsp 458  
QY 1428 CTCCTGAATATTCGATGATCATCTGCATCATGCCAGAAACCACTATGTTCACT 1487  
Db 459 ValThrGluTyrCysAsnGlyThrSerGluValCysGluGluAspPhePheValGlnAsp 478  
QY 1488 GGGCATCCGTGTGAGTGAATCAATGATCTGTATAGATGGAGTTTCTATGAGTGGGAT 1547  
Db 479 GlyHisProCysAlaGluGlnLysTrpIleCysIleAsnGlyThrCysGlnSerGlyAla 498  
QY 1548 AACAATGTACAGACACATTTGGCAAGAAGTAGTGTGGCCCTTCAGAAATGTTATCT 1607  
Db 499 GlnGlnCysArgAspLeuPheGlyThrAspAlaAspTyrGlyThrLysGluCysTyrSer 518  
QY 1608 CACTTAATTCAGAGCTGATGATCTGGAACCTGCTGATAGTATGATTCAGATACACA 1667  
Db 519 GluLeuAsnSerLysSerAspIleSerGlySerCysGlyIleThrProThrGlyTyrLys 538  
QY 1668 CAGTGTGAAGTCAATCTGAGTGGGAAATTAATATGTAATATGATGAGTAAATTT 1727  
Db 539 AspCysAlaProAsnAspArgMetCysGlyLysLeuIleCysIleThrGlnSerGluAsp 558  
QY 1728 TTATTACAAATCCAGAGCCACTTATTATTATGCGCAACATAAGTGGACATCTCTGCATT 1787  
Db 559 IleLeuLysMetArgSerAlaIleValIleTyrAlaAsnIleSerGlyGlnIleCysIle 578  
QY 1788 GCTGTGAATTTGCCAGTATCATCGACAGCAGCAAGATGGGATAAAGATGGAAT 1847  
Db 579 SerLeuGluTyrProProGlyHisLysGluSerLysLysMetCysValArgAspGlyThr 598  
QY 1848 TCTTGTGGTTCAATAAGTTTGCAGGAATCAAGATGTGTGAGTTCTTCATCTGGGT 1907  
Db 599 ValCysGlySerGlyLysValCysLeuAsnGlnGluCysValGluAspThrPheLeuAsn 618  
QY 1908 TATGATTGTACTACTGCAAAATGCAATGATAGAGGTGTATGCAATACAAAGAGCACTGT 1967  
Db 619 TyrAspCysThrProGluLysCysAsnHisGlyValCysAsnAsnLysLysHisCys 638  
QY 1968 CACTGATGCTTCATATTACCTCGAGATGCTGATGCTCAATCAGATCTATGCTGGT 2027  
Db 639 HisCysGluProThrTyrLeuProProAspCysLysAsnThrGluAspThrTrpProGly 658  
QY 2028 GGGAGTATTCAGAGTGGCAATTTTCCACCTGAGCTATACCGCAGACCTCCCTGAAAGG 2087  
Db 659 GlySerValAspSerGlyAsnGlnArgAla-----GluSerIleProAlaArg 675  
QY 2088 CGGTACATTGAGAACTTTTACCATTTCCAAAACCAATGAGATGGCCATTTTCTTATTCA 2147  
Db 676 SerTyrValAlaSerAlaTyrArgSerLysSerAlaArgTrpProPheLeuIleIle 695  
QY 2148 CTTTCTTTTATTTTCTGTGACTGATGCTATATGTTGTAAGTAAAGTTAATTTCCAAAGG 2207

Db 696 ProPheTyrValValIleLeuValLeuIleGlyMetLeuValLysValTyrSerGlnArg 715  
QY 2208 AAAAAATGGAGACTGAGTCACTATTCAAGCGATGAGCAACCTGAAAGTGAAGTGAACCT 2267  
Db 716 LysLysTrpArgMetAspPheSerSerGluGluGlnPheGluSerGluSerGluSer 735  
QY 2268 AAA 2270  
Db 736 Lys 736  
RESULT 9  
AD02\_MOUSE STANDARD; PRT; 735 AA.  
AC Q60718; Q60814; Q9D4G3; Q9QWJ0;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)  
DE (Fertilin beta subunit) (PH-30) (PH30) (PH30-beta).  
GN Name=Adam2; Synonyms=Ftnb;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=96019260; PubMed=7593287;  
RA Evans J.P., Schultz R.M., Kopf G.S.;  
RT "Mouse sperm-egg plasma membrane interactions: analysis of roles of  
RT egg integrins and the mouse sperm homologue of PH-30 (fertilin  
RL beta.");  
RL J. Cell Sci. 108:3267-3278(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Gupta S.K., Alves K., Palladino L.O., Mark G.E., Hollis G.F.;  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikiado I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla S., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Whistled C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L.,  
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume M., Inotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,  
RA Yasunishi A., Yoshino M., Sasaki D., Waterston R., Landers E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE OF 17-735 FROM N.A.



CC TISSUE=Testis;  
 RX MEDLINE=95269891; PubMed=7750654;  
 RA Wolfberg T.G., Straight P.D., Gerena R.L., Huovila A.-P.,  
 RA Primakoff P., Wyles D.G., White J.M.;  
 RT "ADAM, a widely distributed and developmentally regulated gene family  
 RT encoding membrane proteins with a disintegrin and metalloprotease  
 RT domain";  
 RL Dev. Biol. 169:378-383(1995).  
 CC -|- FUNCTION: Sperm surface membrane protein that may be involved in  
 CC sperm-egg plasma membrane adhesion and fusion during  
 CC fertilization. Could have a direct role in sperm-zona binding or  
 CC migration of sperm from the uterus into the oviduct. Interactions  
 CC with egg membrane could be mediated via binding between its  
 CC disintegrin-like domain to one or more integrins receptors on the  
 CC egg. This is a non catalytic metalloprotease-like protein (By  
 CC similarity).  
 CC -|- SUBUNIT: Heterodimer with ADAM1/fertilin alpha.  
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -|- TISSUE SPECIFICITY: Expressed specifically in testis.  
 CC -|- DOMAIN: A tripeptide motif (QDE) within disintegrin-like domain  
 CC could be involved in the binding to egg integrin receptor and thus  
 CC could mediate sperm/egg binding (By similarity).  
 CC -|- PTM: The signal and the metalloprotease domain are cleaved during  
 CC the epididymal maturation of the spermatozoa.  
 CC -|- SIMILARITY: Belongs to peptidase family M12B.  
 CC -|- SIMILARITY: Contains 1 disintegrin domain.  
 CC -|- SIMILARITY: Contains 1 EGF-like domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U16242; AAA90980.1; -;  
 DR EMBL; U38806; RAD04207.1; -;  
 DR EMBL; AK016550; BAB30298.1; -;  
 DR EMBL; U22057; AAA74921.1; -;  
 DR HSSP; P17347; 2ECH.  
 DR MEROPS; M12.950; -;  
 DR MGD; MGI:1340894; Adam2.  
 DR InterPro; IPR006586; ADAM cysteine.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR000742; EGF 2.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR008994; Nucleic acid OB.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR Pfam; PF00200; Disintegrin; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; RepTolysein; 1.  
 DR PRINTS; PR00289; DISINTEGRIN.  
 DR ProDom; PD000664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR PROSITE; PS02015; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS00214; DISINTEGRIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 KW Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
 FT SIGNAL 1 18  
 FT PROPEP 19 180  
 FT CHAIN 181 735  
 FT DOMAIN 19 686  
 FT TRANSMEM 687 707  
 FT DOMAIN 708 735  
 FT DOMAIN 181 388  
 FT DOMAIN 389 479  
 FT DOMAIN 480 611

FT DOMAIN 615 648  
 FT DISULFID 293 376  
 FT DISULFID 335 360  
 FT DISULFID 449 462  
 FT DISULFID 619 630  
 FT DISULFID 624 636  
 FT DISULFID 638 647  
 FT CARBOHYD 128 128  
 FT CARBOHYD 226 226  
 FT CARBOHYD 279 279  
 FT CARBOHYD 359 359  
 FT CARBOHYD 463 463  
 FT CARBOHYD 489 489  
 FT CARBOHYD 569 569  
 FT CARBOHYD 585 585  
 FT CONFLICT 2 2  
 FT CONFLICT 17 20  
 FT CONFLICT 24 25  
 FT CONFLICT 113 113  
 FT CONFLICT 234 242  
 FT CONFLICT 246 247  
 FT CONFLICT 286 286  
 FT CONFLICT 331 332  
 FT CONFLICT 382 382  
 FT CONFLICT 658 658  
 FT CONFLICT 679 679  
 FT CONFLICT 712 712  
 SQ SEQUENCE 735 AA; 82374 MW; 75EC8529CF5B8E2B CRC64;  
 Alignment Scores:  
 Pred. No.: 6,21e-177 Length: 735  
 Score: 2453.50 Matches: 442  
 Percent Similarity: 75.81% Conservative: 119  
 Best Local Similarity: 59.73% Mismatches: 166  
 Query Match: 51.16% Indels: 13  
 DB: 1 Gaps: 8  
 US-10-054-683-18 (1-2640) x AD02\_MOUSE (1-735)  
 QY 72 ATGTGG---GTCTTGTCTCTCAGCGGG-----CTCGCGGGGCTG----- 110  
 DB 1 MetTyrLeuLeuLeuLeuLeuSerGlyLeuSerGluLeuGlyLeuSerGlnSer 20  
 QY 111 CGGATGACAGTAATTTGATAGTTTACCTGTGCAAAATTTACAGTCCGCGAGAAATACGG 170  
 DB 21 GlnThrGluGlyThrArgGluLysLeuHisValGlnValThrValProGluLysIleArg 40  
 QY 171 TCATATAAAGCAAGGAATTTGAATCGCAGGCATCTACAAATTTGTAATTTGAAGGAAA 230  
 DB 41 SerValThrSerAsnGlyTyrGluThrGlnValThrTyrAsnLeuLysIleGluGlyLys 60  
 QY 231 CCATATACTGTGAATTTAATGCAAAAAAATCTTTTACCCCATATTTTAGAGTTTACAGT 290  
 DB 61 ThrTyrThrLeuAspLeuMetGlnLysProPheLeuProAsnPheArgValTyrSer 80  
 QY 291 TATAGTGGCAGCAATTTATGAACACCTTGACCAAGATTTTCAGAAATTTCTGCCACTAC 350  
 DB 81 TyrAspAsnAlaGlyIleMetArgSerLeuGluGlnLysPheGlnAsnIleCysTyrPhe 100  
 QY 351 CAGGGGTATATTGCAAGTTATCCAAATCTGTGCTGATGTTAGTACATCTACTGGACTC 410  
 DB 101 GlnGlyTyrIleGluGlyTyrProAsnSerMetValIleValSerThrCysThrGlyLeu 120  
 QY 411 AGGGCGGTACTACAGTTTGAATTTAGTATTGGAATAGAACCCCTGGAGTCTTCAGTT 470  
 DB 121 ArgGlyPheLeuGlnPheGlyAsnValSerTyrGlyIleGluProLeuGluSerSer 140  
 QY 471 GCGTTTGAACATGTAATTTACCAAGTAAACATAGAACACCATGTTCTCTTATATAT 530  
 DB 141 GlyPheGluHisValIleTyrGlnValGluProGluLysGlyGlyAlaLeuLeuTyrAla 160  
 QY 531 GAGAAGGATATTGAATCAAGAGATCTGCTCTTTAAATTTACAAAGCGCAGACGACGACAA 590  
 EGF-like.  
 By similarity.  
 By similarity.  
 Potential.  
 By similarity.  
 By similarity.  
 N-linked (GlcNAc. .) (Potential).  
 N-linked (GlcNAc. .) (Potential).  
 N-linked (GlcNAc. .) (Potential).  
 N-linked (GlcNAc. .) (Potential).  
 N-linked (GlcNAc. .) (Potential).  
 N-linked (GlcNAc. .) (Potential).  
 N-linked (GlcNAc. .) (Potential).  
 W -> R (in Ref. 2).  
 LSQS -> IRHE (in Ref. 4).  
 GT -> A (in Ref. 4).  
 I -> M (in Ref. 1).  
 LRFWMNDENK -> WNFQWMKQ (in Ref. 4).  
 TG -> QA (in Ref. 4).  
 A -> L (in Ref. 4).  
 DV -> RRL (in Ref. 4).  
 R -> T (in Ref. 2).  
 S -> T (in Ref. 4).  
 A -> R (in Ref. 2).  
 Q -> P (in Ref. 1).  
 75EC8529CF5B8E2B CRC64;











QY 1551 CAATGTACACACATTTGGCAAGAGTAGTGGCTTGGCCCTTCAGAAATGTTATCTTCAC 1610  
 Db 501 GlnCysAlaLysLeuPheGlyLysGlyAlaGlnGlyAlaProPheAlaCysPheLysGlu 520  
 QY 1611 CTTAATCAAGACTGATGATCTGGAACCTGGTATAAGTGAATCAGGATACACACAG 1670  
 Db 521 ValAsnSerLeuHisGluThrSerGluAsnGlyPheLysAsnSerGlnProLeuPro 540  
 QY 1671 TGTGAAGCTGCAATCTGAGTGGGAAATTAATATGTAATATGAGTAAATTTTA 1730  
 Db 541 CysGluArgLysAspValLeuLysGlyLysLeuAlaCysValGlnProHisLysAsnAla 560  
 QY 1731 TTACAATCCAGAGCCATATATTTATGTCACCAATAGTGGACATCTCTGATGCT 1790  
 Db 561 TyrLysSerAspLeuGlnThrValThrSerTyrLeuGlnAspHisValCysValSer 580  
 QY 1791 GTG-----GAATTTGCCAGTGATCATGACAGACAGCCAAAGATGTGATAAAA 1838  
 Db 581 IleAlaThrGlySerSerMetArgSerAspGlyThrAspAsn-----AlaThrValAla 598  
 QY 1839 GATGGAACCTTCTGTGTTCAAAATAGGTTTGCAGGAAATCAAGATGTGTGAGTCTTCA 1898  
 Db 599 AspGlyThrMetCysGlyProGluMetTyrCysValAsnLysThrCysArgLysValHis 618  
 QY 1899 TACTGGTTATGTTGT-----ACTACTGACAAATGCAATGATGAGTGTATGCAATAAC 1955  
 Db 619 LeuThrGlyTyrAsnGlyAsnThrThrLysCysLysGlyLysGlyLysGlyLysGly 638  
 QY 1956 AAAAAGCAGTCACTAGTGTCTCATATTTACCTCCAGTGTCTGATGATCAATCAGAT 2015  
 Db 639 PheGlyAsnGlyLysGlyLysPheProGlyHisLysProAspCysLysPheGlnPheGly 658  
 QY 2016 CTATGGCTGTGGAGTATGACAGTGGCAATTTTCCACCTGTAGTATACACGACGAGA 2075  
 Db 659 Ser---ProGlyGlySerIleAspAspGlyAsnPhe-----Gln 670  
 QY 2076 CTCCTGGAAGGCTACATGAGACATTTACATTCACCAACCAATGAGATGG----- 2129  
 Db 671 LysSerAspGluPheThrGluLysGlyTyrAsnAlaHisTrpAsnAsnTrpPheIle 690  
 QY 2130 ---CCATTTTCTTATTCATCTCTCTTCTTATTTATTTCTGCTACTGATGCTATAATG 2186  
 Db 691 LeuSerPheThrIleValLeuProPheIleIlePheThrIleValIle----- 707  
 QY 2187 GTGAAGTAAATTTCCAAAGGAAAAATGAGGA 2219  
 Db 708 -----PheLysArgAsnGluLeuArg 714

RESULT 12  
 AD18\_HUMAN STANDARD; PRT; 739 AA.  
 AC Q9Y307;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE ADAM 18 precursor (A disintegrin and metalloproteinase domain 18)  
 DE (Transmembrane metalloproteinase-like, disintegrin-like, and cysteine-rich protein III) (tMDC III).  
 GN Name=ADAM18; Synonyms=tMDC3;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA Hall L., Frayne J., Dimsey E.A.;  
 RT "Nucleotide sequence of the human tMDC III sperm surface protein transcript.";  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Sperm surface membrane protein that may be involved in spermatogenesis and fertilization. This is a non catalytic metalloproteinase-like protein (By similarity).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Expressed specifically in testis.  
 CC -1- DOMAIN: A tripeptide motif (ECD) within disintegrin-like domain could be involved in the binding to egg integrin receptor and thus could mediate sperm/egg binding (By similarity).  
 CC -1- PTM: The prodomain and the metalloproteinase-like domain are cleaved during the epididymal maturation of the spermatozoa (By similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin domain.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AJ133004; CAB40812.1; -.  
 CC HSPB; P83469; IMPZ.  
 CC DR MEROPS; M12.957; -.  
 CC DR Genew; HGNC:196; ADAM18.  
 CC DR GO; GO:0005624; C:membrane fraction; TAS.  
 CC DR GO; GO:0008237; F:metalloproteinase activity; TAS.  
 CC DR GO; GO:0007283; P:spermatogenesis; TAS.  
 CC DR InterPro; IPR006586; ADAM cysteine.  
 CC DR InterPro; IPR001762; Disintegrin.  
 CC DR InterPro; IPR006209; EGF-like.  
 CC DR InterPro; IPR001590; Peptidase\_M12B.  
 CC DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 CC DR Pfam; PF00200; Disintegrin; 1.  
 CC DR Pfam; PF01562; Pep M12B\_propep; 1.  
 CC DR Pfam; PF01421; Reprolysin; 1.  
 CC DR ProDom; PD000664; Disintegrin; 1.  
 CC DR SMART; SM00608; ACR; 1.  
 CC DR SMART; SM00505; DISIN; 1.  
 CC DR PROSITE; PS00215; ADAM\_MBP; 1.  
 CC DR PROSITE; PS00437; DISINTEGRIN\_1; FALSE\_NEG.  
 CC DR PROSITE; PS00214; DISINTEGRIN\_2; 1.  
 CC DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 CC DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 CC KW EGF-like domain; Glycoprotein; Signal; Spermatogenesis; Transmembrane.  
 FT SIGNAL 1 16 Potential.  
 FT PROPEP 17 184 Potential.  
 FT CHAIN 185 739 ADAM 18.  
 FT DOMAIN 177 687 Extracellular (Potential).  
 FT TRANSMEM 688 708 Potential.  
 FT DOMAIN 709 739 Cytosolic (Potential).  
 FT DOMAIN 177 389 Metalloproteinase-like.  
 FT DOMAIN 390 479 Disintegrin-like.  
 FT DOMAIN 480 622 Cys-rich.  
 FT DOMAIN 620 654 EGF-like.  
 FT DISULFID 293 376 By similarity.  
 FT DISULFID 335 360 By similarity.  
 FT DISULFID 458 483 Potential.  
 FT DISULFID 624 636 By similarity.  
 FT DISULFID 630 642 By similarity.  
 FT DISULFID 644 653 By similarity.  
 FT CARBOHYD 36 36 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 76 76 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 122 122 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 149 149 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 156 156 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 177 177 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 294 294 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 359 359 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 465 465 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 561 561 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 611 611 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 625 625 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 739 AA; 82856 MW; 2D8BE9A975072CDD CRC64;













```

Db 598 AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg 617
Qy 1899 TACTTG-----GTTTATGTTTACTACTGACAAATGCAATGCAATGATAGAGTGTATGC 1949
Db 618 IleIleIysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys 636
Qy 1950 AATAACAAGAAGCACTCTCACTGAGTGCCTTCATATTACTCCAGATTCTCAGTTCAA 2009
Db 637 AspSerArgAsnLysCysHisCysSerProGlyTyrLysProProAsnCysGlnIleArg 656
Qy 2010 TCAGATCTATGGCTGGTGGAGTATT-----GACAGTGGCAATTTTCCACCT 2057
Db 657 Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle----- 671
Qy 2058 GTAGCTATACAGCCAGACCTCCCTGAAAGGCGGTACATTGAGAACATTTACCATTCCAA 2117
Db 672 -----MetGluArgAlaSerGlyLysThr 679
Qy 2118 CCAATGAGATGG-----CCATTTTCTTATTCTATTCTTTTATTATTCTGTGTA 2171
Db 680 GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal----- 696
Qy 2172 CTGATTCTATAATGTCGAAAGTTAAATTTCCAAAGGAAATGAGAACTGAGGACTAT 2231
Db 697 ---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu--- 714
Qy 2232 TCAAGCATGAGCAACTGAAAGTCAGAGTGAACCTAAAGG 2273
Db 715 -----GluGluPheProSerSerGluSerLysSerGluGly 726

RESULT 15
Q8TC42 PRELIMINARY; PRT; 787 AA.
AC Q8TC42
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE A disintegrin and metalloprotease domain 32.
GN Name=ADAM32;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettner M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley A.C., Trinchard J.W., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; BC026085; AAH26085.1; --
DR HSSP; P30403; IN4Y.
DR MEROPS; M12_960; --
DR GO; GO:0004222; F-metalloendopeptidase activity; IEA.
DR GO; GO:0007229; P-integrin-mediated signaling pathway; IEA.
DR GO; GO:0006508; P-proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR002870; Peptidase M12B_N.
DR Pfam; PF00200; Disintegrin_1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR ProDom; PD00664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS02115; ADAM_MBP; 1.
DR PROSITE; PS02114; DISINTEGRIN_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00026; EGF_3; 1.
KW Integrin; Metalloprotease; Protease.
SQ SEQUENCE 787 AA; 87931 MW; CE0F54A50466B336 CRC64;

Alignment Scores:
Pred. No.: 2,98e-106 Length: 787
Score: 1519.50 Matches: 309
Percent Similarity: 58.89% Conservative: 135
Best Local Similarity: 40.98% Mismatches: 259
Query Match: 31.68% Indels: 51
DB: 2 Gaps: 18

US-10-054-683-18 (1-2640) x Q8TC42 (1-787)
Qy 72 ATGTGGTCTTGTCTTCTGTCAGCGGGCTCGGGGGCTCGGATGAC---AGTAATTTT 128
Db 4 LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe 21
Qy 129 GATAGTTTACCTGCAATACAGTTCGGAGAAATACGGTCAATAATAAGGAAGGA 198
Db 22 GlnAsnSerLeuLeuGlnIleValIleProGluLysIleGlnThrAsnThrAsnAspSer 41
Qy 189 ATTGAA-----TCGAGGCGATCTACAAAATGTAATTCAGAGGAAACCATATAT 239
Db 42 SerGluIleGluTyrGluGlnIleSerTyrIleIleProIleAspGluLysLeuTyrThr 61
Qy 240 GTGAATTTAATGCAAAAAAATTTTACCCCATATAATTTAGAGTTTACAGTTTAGTGGC 299
Db 62 ValHisLeuLysGlnArgTyrPheLeuAlaAspAsnPheMetIleTyrLeuTyrAsn--- 80
Qy 300 ACAGGAATATGAAACCATCTGACCAAGATTTTCAGAAATTTCTGCCACTACCAAGGAT 359
Db 81 GlnGlySerMetAsnThrTyrSerSerAspIleGlnThrGlnCysTyrTyrArgGlyAsn 100
Qy 360 ATTGAAGTTTATCAAAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 419
Db 101 IleGluGlyTyrProAspSerMetValThrLeuSerThrCysSerGlyLeuArgGlyIle 120
Qy 420 CTACAGTTTCAAAATGTTAGTTATGAATAGAACCCCTGGAGCTCTCAGTTGGCTTTGAA 479
Db 121 LeuGlnPheGluAsnValSerTyrGlyIleGluProLeuGluSerAlaValGluPheGln 140
Qy 480 CATGTAATTTACCAAGTAAACATACAGAACGACAGTCTTCTTATATATATAGAGAGAT 539
Db 141 HisValLeuTyrLysLeuLysAsnGluAspAsnAspIleAlaIleAlaIleAspArgGly 160
Qy 540 ATTGAATCAAGA-----GATCTGTCTCTTTAAATTTACAAAGCCGACAGCCACAGCAA 590
Db 161 LeuLysGluGlnProMetAspAsnIlePheIleSerGluLysSerGluProAlaVal 180

```

```
QY 591 GAT-----TTGCAAGTATATAGAAATGCATGTATATAGTTGAAACAAATTTGTATAT 644
Db      |||
QY 181 ProAspLeuPheProLeuTyLeuGluMetHisIleValValAspLysThrLeuTyAsp 200
Db      |||
QY 645 CATATGGGTCATGATCAACCTGTTGCTCGCTCAAAAGCTTTTCAGTTGATTGGATTGACG 704
Db      |||
QY 201 TyrTrpGlySerAspSerMetIleValThrAsnLysValIleGluIleValGlyLeuAla 220
Db      |||
QY 705 AATGCTATTTTGTGTTTCATTTAATATATACAAATTTCTGCTTCATTTGGAGCTTTGGATA 764
Db      |||
QY 221 AsnSerMetPheThrGlnPheLysValIleValLeuSerSerLeuLeuLeuTrpSer 240
Db      |||
QY 765 GATGAAATAAATTTGCAACCTCGGAGAGCTTAATGAGTATTATACACATTTTAAAGA 824
Db      |||
QY 241 AspGluAsnLysIleSerThrValGlyGluAlaAspGluLeuLysPheLeuGlu 260
Db      |||
QY 825 TGGAACACATCTTATCTTGTGTTTACGTCCTCATGATGTGGCATTCTTACCTTGTGTTTACAGA 884
Db      |||
QY 261 TrpLysGlnSerTyLeuAsnLeuArgProHisAspIleAlaTyLeuLeuIleTyMet 280
Db      |||
QY 885 GAAAAGTCAAATATATGTTGTCGCAACCTTTCAAGGGAAGATGTGTGATGCAAACTATGCA 944
Db      |||
QY 281 AspTyProArgTyLeuGlyAlaValPheProGlyThrMetCysIleThrArgTySer 300
Db      |||
QY 945 GGAGGTGTTTCTGCAACCCAGAACCATTAAGTCTGGAATCACTGTGCAGTTATTTAGCT 1004
Db      |||
QY 301 AlaGlyValAlaLeuTyProLysGluIleThrLeuGluAlaPheAlaValIleValThr 320
Db      |||
QY 1005 CAATATTAGCCTTAGTAGTGGGATCATTATGATGACATTAACAAATGCCAGTCTCA 1064
Db      |||
QY 321 GlnMetLeuAlaLeuSerLeuGlyIleSerTyAspAspProLysLysCysGlnCysSer 340
Db      |||
QY 1065 GGAGCTGCTGTCATTTAATGAATCAGAAAGCAATTCATTTACGTGTGTGAAGATCTTTAGT 1124
Db      |||
QY 341 GluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValLysThrPheSer 360
Db      |||
QY 1125 AACTGCAGCTCGAAGACTTTGCACATTTATTTTAAAGCAGAAAGTCCCGAGTCTTTCAC 1184
Db      |||
QY 361 SerCysSerLeuArgSerPheGlnAsnPheIleSerAsnValGlyValLysCysLeuGln 380
Db      |||
QY 1185 AATCAGCCTCGCTAGATCCCTTTTCAACAGCAGCAGTGTGTGTTATGCAAGCTG 1244
Db      |||
QY 381 AsnLysProGlnMetGlnLys---LysSerProLysProValCysGlyAsnGlyArgLeu 399
Db      |||
QY 1245 GAAGCAGGAGAGTGTAGTGTGGGACTGAACAGGATTTGCGCCTTATTGGAGAAACA 1304
Db      |||
QY 400 GluGlyAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer 417
Db      |||
QY 1305 TGCTGTGATATGCCATGTAGATTTAAAGCCGGTTCAAACTGTGCTGAAGGACCATGC 1364
Db      |||
QY 418 CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyLeuLysGlyLeuCys 437
Db      |||
QY 1365 TGGCAAAACTGTCTATTATTATGTCAAAGAAAGATGTGTAGGCT---TCCTTTGAAGAA 1421
Db      |||
QY 438 CysLeuAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu 457
Db      |||
QY 1422 TGGGACCTCCCTGAATATGAATTTGATCATCTGTCATCATGCCAGAAACCACTATGTT 1481
Db      |||
QY 458 CysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyProAspIleThrLeu 477
Db      |||
QY 1482 CAGACTGGGCAATCGGTGTGAGCTGAATCAATGGATCTGTATAGATGGAGTTTGTATGAGT 1541
Db      |||
QY 478 IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyLeuAspGlyAspCysHisAsp 497
Db      |||
QY 1542 GGGGATAAACATGTACAGACATTTGGCAAGAGATGAGTTTGGCCCTTCAGATGT 1601
Db      |||
QY 498 LeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAlaProPheAlaCys 517
Db      |||
QY 1602 TATTCTCACCTTAATCAAGACTGATGTATCTGGAACCTGTGTATAGT---GATTCA 1658
Db      |||
QY 518 TyGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn 537
Db      |||
QY 1659 GGATACACACAGTGTGAAGCTGACAAATCTGCAGTGGCGGAAATTAATATGTAATAATATGTA 1718
```

```
Db      |||
QY 538 LysTyValPheCysGlyTrpArgAsnLeuIleCysGlyArgLeuValCysThrTyPro 557
QY 1719 GGTAAATTTTATTACAAATTTCCAAGAGCCACTATTATTATGCAACATAATAGTGGACAT 1778
Db      |||
QY 558 ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyLeuAlaPheValArgAspSer 577
QY 1779 CTCTGATTTGCTGTGGAATTTGCCAGTGTATCATGACAGACAGCCAAAGATGTGGATTA 1838
Db      |||
QY 578 ValCysIleThrValAspTyLeuLysLeuProArgThrValProAspProLeuAlaValLys 597
QY 1839 GATGGAACCTCTTGTGTTCAATAAGTTTGCAGGAATCAAAAGATGTGTGAGTTCTTCA 1898
Db      |||
QY 598 AsnGlySerGlnCysAspIleGlyArgValAsnArgGluCysValGluSerArg 617
QY 1899 TACTTG-----GGTTATGATTGTACTACTACAAATGCAATGCAATGATAGAGGTATGC 1949
Db      |||
QY 618 IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys 636
QY 1950 AATAACAAAAGCACTGTCACTAGTGTCTCATATTTACTCCAGATTTCTCAGTTCAA 2009
Db      |||
QY 637 AspSerArgAsnLysCysHisCysSerProGlyTyLeuProAsnCysGlnIleArg 656
QY 2010 TCAGATCTATGGCCTGGTGGAGTATT-----GACAGTGGCAATTTTCCACCT 2057
Db      |||
QY 657 Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle----- 671
QY 2058 GTAGCTATACACAGCCAGACTCCCTGAAAGCGCTACATTGAGAAACATTTACCATTCCAA 2117
Db      |||
QY 672 -----MetGluArgAlaSerGlyLysThr 679
QY 2118 CCAATGAGATGG-----CCATTTTCTTATTTCATTCCTTTCTTTATTATTTCGTGTA 2171
Db      |||
QY 680 GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal----- 696
QY 2172 CTGATTGCTATATGTTGAAGTTAATTTCCAAGGAAATAATGGAGAACTGAGGACTAT 2231
Db      |||
QY 697 ---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu--- 714
QY 2232 TCAAGCGATGAGCAACCTGAAAGTGAGAGTGAACTTAAGGG 2273
Db      |||
QY 715 -----GluGluPheProSerSerGluSerLysSerGluGly 726
```

Search completed: January 10, 2005, 22:04:15

Job time : 559.5 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 10, 2005, 21:33:10 ; Search time 66 Seconds  
(without alignments)  
5305.440 Million cell updates/sec

Title: US-10-054-683-18  
Perfect score: 4796  
Sequence: 1 catctgcacttccaactgc.....actgttaccatttcaaaaaa 2640

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 956278

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_p/US10054683/runat\_10012005\_174511\_15520/app\_query\_faasta\_1.2823  
-DB=Issued\_Patents\_AA -QPMF=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10054683 @CGN 1.1.72 @runat\_10012005\_174511\_15520 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3984	83.1	734	2	US-08-765-243-8
2	3984	83.1	734	5	PCT-US95-07295-8
3	3551	74.0	651	1	US-08-264-101-2
4	3551	74.0	651	2	US-08-765-243-2
5	3551	74.0	651	5	PCT-US95-07295-2
6	2440	50.9	735	2	US-08-765-243-6
7	2440	50.9	735	5	PCT-US95-07295-6
8	1512	31.5	457	1	US-08-264-101-4
9	1512	31.5	457	2	US-08-765-243-4
10	1512	31.5	457	5	PCT-US95-07295-4
11	1506.5	31.4	787	4	US-10-000-489-70
12	1493.5	31.1	715	4	US-10-140-002-116

Sequence 204, Appl  
Sequence 2, Appli  
Sequence 15, Appl  
Sequence 32, Appl  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 6, Appli  
Sequence 12, Appl  
Sequence 4, Appli  
Sequence 5, Appli  
Sequence 10, Appl  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 18, Appl  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 6, Appli  
Sequence 3, Appli  
Sequence 14, Appl  
Sequence 4, Appli  
Sequence 1, Appli

Sequence 204, Appl  
Sequence 2, Appli  
Sequence 15, Appl  
Sequence 32, Appl  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 6, Appli  
Sequence 12, Appl  
Sequence 4, Appli  
Sequence 5, Appli  
Sequence 10, Appl  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 18, Appl  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 6, Appli  
Sequence 3, Appli  
Sequence 14, Appl  
Sequence 4, Appli  
Sequence 1, Appli

#### ALIGNMENTS

RESULT 1  
US-08-765-243-8  
Sequence 8, Application US/08765243  
Patent No. 5935578  
GENERAL INFORMATION:  
APPLICANT: ALVES, KENNETH  
APPLICANT: GUPTA, SUNIL K.  
TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MARY A. APPOLLINA  
STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE  
CITY: RAHWAY  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/765,243  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: APPOLLINA, MARY A  
REGISTRATION NUMBER: 34,087  
REFERENCE/DOCKET NUMBER: 19244Y  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908)594-3462  
TELEFAX: (908)594-4720  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 734 amino acids

; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-765-243-8

## Alignment Scores:

Pred. No.: 0 Length: 734  
 Score: 3984.00 Matches: 734  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 83.07% Indels: 0  
 DB: 2 Gaps: 0

US-10-054-683-18 (1-2640) x US-08-765-243-8 (1-734)

QY 72 ATGTGGTCTGTTCTGCTCAGGGCTGGGGCTGGGATGGACAGTATTTTGAT 131  
 DB 1 MetTrrpValLeuPheLeuSerGlyLeuGlyLeuArgMetAspSerAsnPheAsp 20  
 QY 132 AGTTTACCTGTGCAATTTACAGTTCCGAGAAATACGGTCAATAATAAGGAAGGAATT 191  
 DB 21 SerLeuProValGlnIleThrValProGluLysIleArgSerIleIleLysGluGlyIle 40  
 QY 192 GAATCGCAGCATCTCTCAAAATTTGTAATTGAAGGAAACCATATCTGTGAATTTAATG 251  
 DB 41 GluSerGlnAlaSerTyrLysIleValIleGluGlyLysProTyrThrValAsnLeuMet 60  
 QY 252 CAAAAAATCTTTTACCCCATATTTTACAGTTTACAGTTTACAGTTTACAGTTTATG 311  
 DB 61 GlnLysAsnPheLeuProHisAsnPheArgValTyrSerTyrSerGlyThrGlyIleMet 80  
 QY 312 AAACCACTTCACCAAGATTTTTCAGAAATTTCTGCCACTACCAAGGTTATTTGAAGTTAT 371  
 DB 81 LysProLeuAspGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGlyTyr 100  
 QY 372 CCAAAATCTGTGTGATGTTAGCATGTACTCGACTCAGGGCGTACTACAGTTTGAA 431  
 DB 101 ProlLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGlu 120  
 QY 432 AATGTTAGTTATGGAATAGAACCTCGAGTCTTCAGTTTTCAGTTTTCAGTTTATG 491  
 DB 121 AsnValSerTyrGlyIleGluProLeuGluSerSerValGlyPheGluHisValIleTyr 140  
 QY 492 CAAGTAAACATAGAAAGCAGATGTTTCTCTTATATATAGAGAGATTTTGAATCAAGA 551  
 DB 141 GlnValLysHisLysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSerArg 160  
 QY 552 GATCTGTCTTTAAATTAACAGCGCAGCCACAGCAAGATTTTGCAGATATATAGAA 611  
 DB 161 AspLeuSerPheLysLeuGlnSerAlaGluProGlnGlnAspPheAlaLysTyrIleGlu 180  
 QY 612 ATGCATCTTATAGTTGAAACCAATTTGTAATCATATGAGTCTGATACAACTGTTGTC 671  
 DB 181 MethisValIleValGluLysGlnLeuTyrAsnHisMetGlySerAspThrThrValVal 200  
 QY 672 GCTCAAAAGATTTTCCAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 731  
 DB 201 AlaGlnLysValPheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsnIle 220  
 QY 732 ACAATTAATCTGCTTTCATTTGAGCTTTGGATAGATGAAATAAATAATTTGCACCTGGA 791  
 DB 221 ThrIleIleLeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThrThrGly 240  
 QY 792 GAAGCTAATGATTTATACACATTTTAAAGATGGAACATCTTATCTTGTGTTTACGT 851  
 DB 241 GluAlaAsnGluLeuLeuHisThrPheLeuArgTrrpLysThrSerTyrLeuValLeuArg 260  
 QY 852 CCTCATGATGTCGCAATTTTACTTTGTTTACAGAGAAAGTCAAAATTTGTTGTCGAACC 911  
 DB 261 ProHisAspValAlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAlaThr 280  
 QY 912 TTTCAGGGAAGATGTGTGATGCAAACTATCGCAGAGGTGTTGTTCTGCCCCGAGAAC 971

DB 281 PheGlnGlyLysMetCysAspAlaAsnTyrAlaGlyValValLeuHisProArgThr 300  
 QY 972 ATAAAGTCTGAATCACTTGCAGTTATTTTAGCTCAATTTATTTAGCTTCAATTTAGTATGGGATC 1031  
 DB 301 IleSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGlyIle 320  
 QY 1032 ACTTATGATGACATTAACAAATCCAGTCTCAGGAGCTGTCTGCAATTTATGAATCCAGAA 1091  
 DB 321 ThrTyrAspAspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnProGlu 340  
 QY 1092 GCATTTTCACTTCACTGCTGTGAAGATCTTTAGTAACTGCGACTTCAAGACTTTGCACAT 1151  
 DB 341 AlaIleHisPheSerGlyValLysIlePheSerAsnCysSerPheGluAspPheAlaHis 360  
 QY 1152 TTTATTTTCAAGCAGAAAGTCCAGTCTCTTCAAAATCAGCTCGCTTAGATCTCTTTTTC 1211  
 DB 361 PheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhePhe 380  
 QY 1212 AAACACAGCAGTGTGTGTATGCAAGCTGGAAGCAGGAGGAGGTGTGTGTGGG 1271  
 DB 381 LysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluCysAspCysGly 400  
 QY 1272 ACTGAACAGGATGTGCTCTTATTCAGAAACATGCTGTGATATTCACCATCTAGATTT 1331  
 DB 401 ThrGluGlnAspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPhe 420  
 QY 1332 AAAGCCGGTTCAAACTGTCTGAAGACCATGTGCAAAACCTGTCTTATTTATGTCAAAA 1391  
 DB 421 LysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLys 440  
 QY 1392 GAAAGAAATGTGTAGGCTTCTTTGAAGAAATGCGACTCTCCTGAATATTTGCAATGATCA 1451  
 DB 441 GluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGlySer 460  
 QY 1452 TCTGCACTCATGCCAGAAACCACTATGTCAGCTGGGCATCCGTGTGACTGAATCAA 1511  
 DB 461 SerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGln 480  
 QY 1512 TGATCTGTATAGATGGAGTTTGTATGAGTGGGGATAAAACAATGTACAGACACATTTGGC 1571  
 DB 481 TrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGly 500  
 QY 1572 AAAGAAGTGTAGTTTGGCTTCCAGAAATGTTATTTCTACCTTAATTTCAAAGACTGATGA 1631  
 DB 501 LysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspVal 520  
 QY 1632 TCTGGAAACTGTGATTAAGTATTTCAGATATACACAGTGTCAAGTGTCAATCTGCAG 1691  
 DB 521 SerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGln 540  
 QY 1692 TCGGGAATAATATATGTAATATGTAAGTAAATTTTATTTATCAAAATTTCCAAGAGCCACT 1751  
 DB 541 CysGlyLysLeuIleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAlaThr 560  
 QY 1752 ATTATTTATGCCAATAAGTGGACATCTCTGCAATTCCTGTGGAATTTGCCAGTATCAT 1811  
 DB 561 IleIleTyrAlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHis 580  
 QY 1812 GCAGACAGCCAAAGATGTGATAAAGATGGAACCTTCTTGTGTTTCAAAATAAGTTTGC 1871  
 DB 581 AlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCys 600  
 QY 1872 AGGAATCAAAAGATGTGTGAGTCTTCTCATACTTCGGTTATGATTTGTTACTACTGACAAATGC 1931  
 DB 601 ArgAsnGlnArgCysValSerSerSerTyrLeuGlyTyrAspCysThrThrAspLysCys 620  
 QY 1932 AATGATAGAGGTGTATGCAATAACAAAGACCTGTCACCTGTAGTGTGTTTACCT 1991  
 DB 621 AsnAspArgGlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeuPro 640  
 QY 1992 CCAGATTGCTCAGTTCAATCAGATCTATGCGCTGGTGGGAGTATTGACAGTGGCAATTTT 2051  
 DB 641 ProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsnPhe 660



1272 ACTGAACAGGATTGTGCCCTTATTGGAGAAACATGCTGTGATATTTGCCATGTAGATTTT 1331  
Db ThrGluGlnAspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPhe 420  
1332 AAAGCCGGTCAAACTGTGTGAAGGACCATGCTGCGAAACATGCTCTATTATGTCAAAA 1391  
Db LysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLys 440  
1392 GAAAGATGTAGTCCCTTCTTTGAAGAATGCCACCTCCCTGAATATTCGAATGGATCA 1451  
Db GluArgMetCysArgProSerPheGluGluCysAspLeuProGluThrCysAsnGlySer 460  
1452 TCTGCATCATGCCAGAAAACCATATGTTCCAGACTGGGCATCCGTGTGCACTGAATCAA 1511  
Db SerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGln 480  
1512 TGATCTGTATAGATGAGTTTGTATGAGTGGGATAAACATGTACAGACACATTTGGC 1571  
Db TrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGly 500  
1572 AAAGAAGTGTAGTTTGGCCCTTCAAGATGTTTATCTCACCTTAATTCAAAGACTGATGA 1631  
Db LysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspVal 520  
1632 TCTGGAACCTGTGTATAGTATTTCAGGATACACAGATGTGAAGCTGACAAATCTGCAG 1691  
Db SerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGln 540  
1692 TGGGAAAATTAATATGTAATATGTTAGTTAAATTTTATTAACAATTCGAAGCCACT 1751  
Db CysGlyLysLeuIleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAlaThr 560  
1752 ATTATTTATGCCAATAAGTGACATCTCTGCATTCGTGGAAATTTGCCAGTGATCAT 1811  
Db IleIleTyrAlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHis 580  
1812 GCAGACAGCCAAAGATGTGGATAAAAGATGGAATCTTCTGTGGTTCCAAATAAGGTTTC 1871  
Db AlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCys 600  
1872 AGGATCAAGATGTGAGTTCTTCTACTATGGTTATGATTTGATCTACTACTGACAAATGC 1931  
Db ArgAsnGlnArgCysValSerSerTyrLeuGlyTyrAspCysThrThrAspLysCys 620  
1932 AATGATAGAGTGATGCAATAACAAAAGACATGTCACCTGCTAGTGTCTTCATATTTACCT 1991  
Db AsnAspArgGlyValCysAsnAsnLysHisCysHisCysSerAlaSerTyrLeuPro 640  
1992 CCAGATTGCTCAGTTCAATCAGATCTATGGCTGTGGAGTATTTGACAGTGGCAATTTT 2051  
Db ProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsnPhe 660  
2052 CCACCTGTAGTATACAGCCAGACTCCCTGAAAGGCGCTACATTTGAGAACATTTACAT 2111  
Db ProProValAlaIleProAlaArgLeuProGluArgTyrIleGluAsnIleTyrHis 680  
2112 TCCAAACCAATGATGAGTGGCAATTTCTTATTCATCTCTTCTTATTTATTTCTGGTA 2171  
Db SerLysProMetArgTrpProPhePheLeuPheIlePheCysVal 700  
2172 CTGATTCCTATAATGGTGAAGTTTAATTTCCAAAGGAAAAAATGAGAACTGAGGACTAT 2231  
Db LeuIleAlaIleMetValLysValAsnPheGlnArgLysLysTyrAspThrGluAspTyr 720  
2232 TCAAGCCATGAGCACTGAAAGTGAGAGTGAACTAAAGG 2273  
Db SerSerAspGluGlnProGluSerGluSerGluProLysGly 734

RESULT 3

US-08-264-101-2

; Sequence 2, Application US/08264101

; Patent No. 5693496

GENERAL INFORMATION:  
APPLICANT: ALVES, KENNETH  
APPLICANT: GUPTA, SUNIL K.  
APPLICANT: HOLLIS, GREGORY F.  
TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MARY A. APPOLLINA  
STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE  
CITY: RAHWAY  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/264,101  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: APPOLLINA, MARY A.  
REGISTRATION NUMBER: 34,087  
REFERENCE/DOCKET NUMBER: 19244  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908)594-3462  
TELEFAX: (908)594-4720  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 651 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-264-101-2

## Alignment Scores:

Pred. No.: 0 Length: 651  
Score: 3551.00 Matches: 650  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 74.04% Indels: 0  
DB: 1 Gaps: 0

US-10-054-683-18 (1-2640) x US-08-264-101-2 (1-651)

Qy 324 CAAGATTTTCAGAAATTTTCGCCACTACCAAGGTTATTTGAAGTTATCCAAATCTGTG 383  
Db 2 GlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGlyTyrProLysSerVal 21  
Qy 384 GTCATGTGTAGCAGTACTGACCTCAGGGCGTACTACAGTTTGAAATGTTAGTTAT 443  
Db 22 ValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGluAsnValSerTyr 41  
Qy 444 GGAATAGAACCCCTGGAGTCTTCAGTTGGCTTTTCAACATCTAATTTTACCAAGTAAACAT 503  
Db 42 GlyLeuProLeuGluSerSerValGlyPheGluHisValIleTyrGlnValLysHis 61  
Qy 504 AAGAAAGCAGATGTTCTTATATATATGAGAAGATATTGAATCAAGAGATCTGTCTTT 563  
Db 62 LysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSerArgAspLeuSerPhe 81  
Qy 564 AAATTACAAGCCAGACAGCCACAGAGATTTTGCAGATATATAGAAATGCATGTTATA 623  
Db 82 LysLeuGlnSerAlaGluProGlnGlnAspPheAlaLysTyrIleGluMetHisValIle 101  
Qy 624 GTTCAAAAACAATTTGATATATCATATGGGCTCTGATACAACTGTTGTCTCAAAAAGTT 683  
Db 102 ValGluLysGlnLeuTyrAsnHisMetGlySerAspThrThrValValAlaGlnLysVal 121  
Qy 684 TTCAGTTGATTGGATTGACGAATGCTATTTTGTTCATTATTAATATATATTTCTG 743



Db 122 PheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsnIleThrIleIleLeu 141  
Qy 744 TCATTGAGGAGCTTTGGATAGATGAAATAAATTCGAACCTCGGAGAGCTAAATCAG 803  
Db 142 SerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThrThrGlyGluAlaAsnGlu 161  
Qy 804 TTATTACACACATTTTAAAGATGAAACATCTTATCTTGTTTACGTCTCCTCATGATGTG 863  
Db 162 LeuLeuHisThrPheLeuArgTrpLysThrSerTyrLeuValLeuArgProHisAspVal 181  
Qy 864 GCATTTTACTGTTTACAGAGAAAGTCAAATATGTTGTGCAACCTTTCAGAGGGAAG 923  
Db 182 AlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAlaThrPheGlnGlyLys 201  
Qy 924 ATGTGTGATGCAAACTATGAGGAGGCTGTTGTTCTGCACCCAGAACCATAGTCTCGAA 983  
Db 202 MetCysAspAlaAsnTyrAlaGlyValValLeuHisProArgThrIleSerLeuGlu 221  
Qy 984 TCATTGAGTATTTAGCTCAATATTATGAGCCCTTAGTATGGGATCACTTATGATGAC 1043  
Db 222 SerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGlyIleThrTyrAspAsp 241  
Qy 1044 ATTAACAAATGCCAGTCTCAGGAGCTGCTGCATTATGATCCAGAGCAATTCATTTC 1103  
Db 242 IleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnProGluAlaIleHisPhe 261  
Qy 1104 AGTGGTGTGAAGATCTTTAGTAACCTGACAGCTTCCAGACCTTTGCACATTTTATTTCAAAG 1163  
Db 262 SerGlyValLysIlePheSerAsnCysSerPheGluAspPheAlaHisPheIleSerLys 281  
Qy 1164 CAGAGTCCAGTGTCTTCAATACAGCTCGCTTAGATCCTTTTTCACACAGCAAGCA 1223  
Db 282 GlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhePheLysGlnGlnAla 301  
Qy 1224 GTGTGTGTAATGCAAGCTGAGAGGAGGAGAGTGTGACTGTGGAGCTGACAGAT 1283  
Db 302 ValCysGlyAsnAlaLysLeuGluAlaGlyGluCysAspCysGlyThrGluGlnAsp 321  
Qy 1284 TGTGCCCTTATTGAGAGAAATGCTGTGATATTGCCACATGTAGATTTAAAGCCGGTTCA 1343  
Db 322 CysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPheLysAlaGlySer 341  
Qy 1344 AACTGTCTGAAGACCATCTCGGAAACTGCTATTATGTCAAAAGAAAGAAATGTGT 1403  
Db 342 AsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLysGluArgMetCys 361  
Qy 1404 AGGCTTTCCTTTGAGAGATGGACCTCCCTGAATATTGCAATGGATCATCTGCATCATGC 1463  
Db 362 ArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGlySerSerAlaSerCys 381  
Qy 1464 CCAGAAACCACTATGTTACAGCTGGGCATCCGTTGGACTGAATCAATGATCTGTATA 1523  
Db 382 ProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGlnTrpIleCysIle 401  
Qy 1524 GATGAGCTTTGTATGAGTGGGATAAACAAATGTACAGACACATTTGGCAAGAGTAGAG 1583  
Db 402 AspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGlyLysGluValGlu 421  
Qy 1584 TTTGGCCCTTCAGAAATGTTATCTCACCTTAATCAAGACTGATGTATCTGGAACCTGT 1643  
Db 422 PheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspValSerGlyAsnCys 441  
Qy 1644 GGTATAGTATTGATTCAGATACACAGCTGTGAACCTGCACAACTGCAAGTCGGGAAATTA 1703  
Db 442 GlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGlnCysGlyLysLeu 461  
Qy 1704 ATATGTAATATGATGAGTAAATTTTATTACAAATTCAGAGCCATTTATTTATGCC 1763  
Db 462 IleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAlaThrIleIleTyrAla 481  
Qy 1764 AACATAAGTGACATCTCTGATGCTGTGGAATTTGCCAGTGTATCATGACAGCCAA 1823

Db 482 AsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHisAlaAspSerGln 501  
Qy 1824 AAGATGTGGATAAAGATGGAACCTCTTGTGGTTCAATTAAGGTTTGAGGAATCAAGA 1883  
Db 502 LysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCysArgAsnGlnArg 521  
Qy 1884 TGTGTGAGTTCTTCATACCTTGGGTATGATTGATCTACTACTGACAAATGCAATGATAGAGT 1943  
Db 522 CysValSerSerTyrLeuGlyTyrAspCysThrThrAspLysCysAsnAspArgGly 541  
Qy 1944 GTATGCAATTAACAAAAGCACTGTCACTGTAGTGTCTTCATATTTACCTCCAGATTGCTCA 2003  
Db 542 ValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeuProProAspCysSer 561  
Qy 2004 GTTCAATCAGATCTATGCGCTGTGGAGTATTGACAGTGGCAATTTTCCACCTGTAGCT 2063  
Db 562 ValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsnPheProProValAla 581  
Qy 2064 ATACCAGCAGACTCCCTGAAAGCGCTACATTGAGAACATTTTACCATTCCAAACCAATG 2123  
Db 582 IleProAlaArgLeuProGluArgTyrIleGluAsnIleTyrHisSerLysProMet 601  
Qy 2124 AGATGGCCATTTTCTTATTCACTTCTTCTTATTATTCTGTGTACTGATTGCTATA 2183  
Db 602 ArgTrpProPhePheLeuPheIleProPhePheIlePheCysValLeuIleAlaIle 621  
Qy 2184 ATGCTGAAAGTTAATTTCCAAAGGAAAAATGGAGAACTGAGGACTATTCAAGCGATGAG 2243  
Db 622 MetValLysValAsnPheGlnArgLysLeuTrpArgThrGluAspTyrSerSerAspGlu 641  
Qy 2244 CAACCTGAAAGTGAGAGTGAACTTAAGGG 2273  
Db 642 GlnProGluSerGluSerGluProLysGly 651

RESULT 4  
US-08-765-243-2  
; Sequence 2, Application US/08765243  
; Patent No. 5935578  
; GENERAL INFORMATION:  
; APPLICANT: ALVES, KENNETH  
; APPLICANT: GUPTA, SUNIL K.  
; APPLICANT: HOLLIS, GREGORY F.  
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MARY A. APOLLINA  
; STREET: P. O. BOX 2000, 126 E. LINCOLN AVENUE  
; CITY: RAHWAY  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/765,243  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: APOLLINA, MARY A  
; REGISTRATION NUMBER: 34,087  
; REFERENCE/DOCKET NUMBER: 19244Y  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908)594-3462  
; TELEFAX: (908)594-4720  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 651 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-765-243-2

## Alignment Scores:

Pred. No.:	0	Length:	651
Score:	3551.00	Matches:	650
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	74.04%	Indels:	0
DB:	2	Gaps:	0

US-10-054-683-18 (1-2640) x US-08-765-243-2 (1-651)

QY	324	CAAGATTTCAGAAATTTCTGCACTACCAAGGGTATATTGAAGGTATCCAAAATCTGTG	383
DB	2	GlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGlyTyrProLysSerVal	21
QY	384	GTGATGTTAGCAGCATGTACTGAGACTCAGGGCGTACTACAGTTTGAATAATGTTAGTTAT	443
DB	22	ValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGluAsnValSerTyr	41
QY	444	GGATAGAACCCCTGGAGTCTTCAGTTGGCTTGAACATGTAATTTACCAAGTAAACAT	503
DB	42	GlyIleGluProLeuGluSerValGlyPheGluHisValIleTyrGlnValLysHis	61
QY	504	AAGAAAGCAGATGTTCTTATATATAATGAGAAGGATATTGAATCAAGAGATCTCTCCTTT	563
DB	62	LysLysAlaAspValSerLeuTyrAsnGlnLysAspIleGluSerArgAspLeuSerPhe	81
QY	564	AAATTACAAAGCGCAGAGCCACAGCAAGATTTTGCAAAGTATATAGAAATGCAATGTTATA	623
DB	82	LysLeuGlnSerAlaGluProGlnGlnAspPheAlaLysTyrIleGluMetHisValIle	101
QY	624	GTTGAAAACAAATGTATATCATATGGGCTCATACACTGTTGCGCTCAAAAAGTT	683
DB	102	ValGlnLysGlnLeuTyrAsnHisMetGlySerAspThrThrValAlaGlnLysVal	121
QY	684	TTCCAGTTGATTTGAGTACCAAAATGCTATTTTCTTCAATTAATATATACAAATATTCTG	743
DB	122	PheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsnIleThrIleLeu	141
QY	744	TCATTGAGGAGCTTTGGATAGATGAAATAAATATGCAACCATCGAGAGCTAATGAG	803
DB	142	SerSerLeuGluLeuThrIleAspGluAsnLysIleAlaThrThrGlyGluAlaAsnGlu	161
QY	804	TTATTACACATTTTAAAGATGAAACATCTTATCTGTTTACCTTCATCATGATGTG	863
DB	162	LeuLeuHisThrPheLeuA-gtrPluSerTyrLeuValLeuArgProHisAspVal	181
QY	864	GCAATTTTACTTTTACAGAAAAAGTCAAAATATGTTGGTGCACACTTTTCAAGGGAAG	923
DB	182	AlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAlaThrPheGlnGlyLys	201
QY	924	ATGTGTGATGAAACTATGAGAGAGGTGTTGTTCTGCACCCAGAACCAATAAGTCTGAA	983
DB	202	MetCysAspAlaAsnTyrAlaGlyValValLeuHisProArgThrIleSerLeuGlu	221
QY	984	TCATTCGAGTTATTTAGCTCAATATTGAGCCTTAGTATGGGATCACTTATGATGAC	1043
DB	222	SerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGlyIleThrTyrAspAsp	241
QY	1044	ATTAACAAATGCCAGTCTCTCAGGAGCTGCTGCAATTAATGATCCAGAACCAATTCATTTC	1103
DB	242	IleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnProGluAlaIleHisPhe	261
QY	1104	AGTGGTGTGAAGATCTTTAGTAACTGACGCTTCGAGACTTTGACATTTTATTTCAAAG	1163
DB	262	SerGlyValLysIlePheSerAsnCysSerPheGluAspPheAlaHisPheIleSerLys	281
QY	1164	CAGAAAGTCCAGTCTCTTCAAAATCAGCCTCGCTTAGATCTCTTTTCAACACGACGCA	1223
DB	282	GlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhePheLysGlnGlnAla	301

RESULT 5

QY	1224	GTGTGTGTTAATGCAAAAGCTGGAAGCAGAGAGAGAGTGTGACTGTGGGACTGAACAGGAT	1283
DB	302	ValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCysGlyThrGluGlnAsp	321
QY	1284	TGTGCCCTTATTGGAGAAACATGCTGTGATATTGCCACATGTAGATTAAAGCGGTTCA	1343
DB	322	CysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPheLysAlaGlySer	341
QY	1344	AACGTGCTCAAGGACCATGCTCGGAAAACCTGCTATTATTTATGTCAAAAGAAAGAAATGCT	1403
DB	342	AsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLysGluArgMetCys	361
QY	1404	AGGCCTTCCCTTTGAAGAATCGGACCTCCCTGAATATTGCAATGGATCATCTGCATCATGC	1463
DB	362	ArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGlySerSerAlaSerCys	381
QY	1464	CGAGAAACCACTATGTTTACAGACTGGGCATCCGTGTGGACTGAATCAATGATCTGTATA	1523
DB	382	ProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGlnTrpIleCysIle	401
QY	1524	GATGGAGTTTGTATGATGGGGATAAACAATGTACAGACACATTTGGCAAGAAGTAGAG	1583
DB	402	AspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGlyLysGluValGlu	421
QY	1584	TTTGGCCCTTCAGAAATGTTATTCTCACCTTAATTTCAAGACTGATGTATCTGGAACCTGT	1643
DB	422	PheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspValSerGlyAsnCys	441
QY	1644	GGTATAGTGAATTCAGGATACACACAGTGTGAAGCTGACAATCTGCAGTCCGGAATAATTA	1703
DB	442	GlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGlnCysGlyLysLeu	461
QY	1704	ATATGTAATATCTAGTAAATTTTATTACAAATTTCCAGAGCCACTATTATTATTATGCC	1763
DB	462	IleCysLysTyrValGlyLysPheLeuGlnIleProArgAlaThrIleIleTyrAla	481
QY	1764	AACATAAGTGACATCTCTGCATTGCTGTGGAATTTGCCAGTGTATCATGACAGACAGCAA	1823
DB	482	AsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHisAlaAspSerGln	501
QY	1824	AGATGTGGATAAAGATGGAATCTTCTGTGGTTTCAATTAAGTTTGCAGGAATCAAGA	1883
DB	502	LysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCysArgAsnGlnArg	521
QY	1884	TGTGTGAGTCTTCTCATACTTGGGTATTGATTGTACTACTCACAATGCAATGATAGAGGT	1943
DB	522	CysValSerSerSerTyrLeuGlyTyrAspCysThrThrAspLysCysAsnAspArgGly	541
QY	1944	GTATGCAATAACAAAAGCACTGCTCACTGAGTGTCTTCATATTATTACCTCCAGATTGCTCA	2003
DB	542	ValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeuProProAspCysSer	561
QY	2004	GTTCAATCAGATCTATGGCTGTGGAGTATTGACAGTGGCAATTTTCCACCTGTAGCT	2063
DB	562	ValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsnPheProProValAla	581
QY	2064	ATACCAAGCAGACTCCCTGAAAGCGCTACATTGAGAACATTTTACCAATTTCCAAACCAATG	2123
DB	582	IleProAlaArgLeuProGluArgTyrIleGluAsnIleTyrHisSerLysProMet	601
QY	2124	AGATGGCCATTTTCTTATTTCATTTCTCTTCTTTTATTATTATTCTGTGACTGATGCTATA	2183
DB	602	ArgTrpProPhePheLeuPheIleProPhePheIleIlePheCysValLeuIleAlaIle	621
QY	2184	ATGCTGAAGTTAATTTCCAAAGAAAATATGAGAACTGAGACACTATTTCACGGATGAG	2243
DB	622	MetValLysValAsnPheGlnArgLysLysTrpArgThrGluAspTyrSerSerAspGlu	641
QY	2244	CAACCTGAAAGTGAGAGTGAACCTTAAAGGG	2273
DB	642	GlnProGluSerGluSerGluProLysGly	651

PCT-US95-07295-2  
; Sequence 2, Application PC/TUS9507295  
; GENERAL INFORMATION:  
; APPLICANT: ALVES, KENNETH  
; APPLICANT: GUPTA, SUNIL K.  
; APPLICANT: HOLLIS, GREGORY F.  
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MARY A. APOLLINA  
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE  
; CITY: RAHWAY  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/07295  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: APOLLINA, MARY A  
; REGISTRATION NUMBER: 34,087  
; REFERENCE/DOCKET NUMBER: 19244Y  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908)594-3462  
; TELEFAX: (908)594-4720  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 651 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-07295-2

Alignment Scores:  
Pred. No.: 0 Length: 651  
Score: 3551.00 Matches: 650  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 74.04% Indels: 0  
DB: 5 Gaps: 0

US-10-054-683-18 (1-2640) x PCT-US95-07295-2 (1-651)

QY	324	CAAGATTTTTCAGAAATTTCTCCCACTACCAAGGGTATATTCAGAGTTATCCAAATCTGTG	383
Db	2	GlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGlyTyrProLysSerVal	21
QY	384	GTGATGGTTAGACATGTACTGGACTCAGGGGGGTACTACAGATTTGAAAATGTTAGTTAT	443
Db	22	ValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGluAsnValSerTyr	41
QY	444	GGATAGAACCCCTGGAGTCTTCAGTTGGCTTTGCAACATGTAATTTACCAAGTAAACAT	503
Db	42	GlyIleGluProLeuGluSerSerValGlyPheGluHisValIleTyrGlnValLysHis	61
QY	504	AAGAAGCAGATGTTCTTATATATGAGAAGGATATTCGAATCAAGAGATCTGCTCTTT	563
Db	62	LysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSerArgAspLeuSerPhe	81
QY	564	AAATTACAAAGCGCAGACGACAGCAAGATTTTGCAAGATATATAGAAATGCATGTTATA	623
Db	82	LysLeuGlnSerAlaGluProGlnGlnAspPheAlaLysTyrIleGluMetHisValIle	101
QY	624	GTTGAAAAACAATGTATATCATATGGGCTGTCATACAACTGTGTGCTCAAAAAGTT	683
Db	102	ValGluLysGlnLeuTyrAsnHisMetGlySerAspThrThrValValAlaGlnLysVal	121

QY	684	TTCCAGTTGATTGGATTGACGAATGCTATTTTGTGTTTCATTATATATACAAATTTCTGTG	743
Db	122	PheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsnIleThrIleLeuLeu	141
QY	744	TCTTCATTGGAGCTTTGGATAGATAGAAATAAAATTCGAACCACTGGAGAGCTAATGAG	803
Db	142	SerSerLeuGluLeuTyrIleAspGluAsnLysIleAlaThrThrGlyGluAlaAsnGlu	161
QY	804	TTATTACACACATTTTAAAGTGAACATCTTATCTGTTTTCACCTCTCATGATGTG	863
Db	162	LeuLeuHisThrPheLeuArgTyrPheThrSerTyrLeuValLeuArgProHisAspVal	181
QY	864	GCATTTTACTTGTTCACAGAGAAAAGTCAAAATATGTTGTTGCAACCTTTTCAAGGGAAG	923
Db	182	AlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAlaThrPheGlnGlyLys	201
QY	924	ATGTGTGATGCAAACTATGACAGAGGTGTTGTTCTGCACCCCAAGAACATAAGTCTCGAA	983
Db	202	MetCysAspAlaAsnTyrAlaGlyGlyValValLeuHisProArgThrIleSerLeuGlu	221
QY	984	TCACTTGCAGTTATTTAGCTCAATATTGAGCCTTAGTATGGGATCACATTATGATGAC	1043
Db	222	SerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGlyIleThrTyrAspAsp	241
QY	1044	ATTAACAAATGCCAGTGCTCAGGAGCTGCTGCATTATGAATCCAGAAGCAATTCATTTC	1103
Db	242	IleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnProGluAlaIleHisPhe	261
QY	1104	AGTGTGTGAAGATCTTTAGTAACCTGCAGCTTCGAAGACTTTGACACATTTTATTTCAAAG	1163
Db	262	SerGlyValLysIlePheSerAsnCysSerPheGluAspPheAlaHisPheIleSerLys	281
QY	1164	CAGAAGTCCCAGTGCTTCAACAATCAGCCTCGCTTAGATCCTTTTCAAACAGCAAGCA	1223
Db	282	GlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhePheLysGlnGlnAla	301
QY	1224	GTGTGTGGTAAATGCAAGCTTGGAAAGCAGAGAGAGTGTGACTGTGGAGCTGAACAGAT	1283
Db	302	ValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCysGlyThrGluGlnAsp	321
QY	1284	TGTGCCCTTATTGAGAAACATGCTGTGATATTTGCCACATGTAGATTTAAAGCGGTTC	1343
Db	322	CysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPheLysAlaGlySer	341
QY	1344	AACTGTGCTGAAGCACCATGCTCGGAAACTGCTATTATGTCARAAGAAAGATCTGT	1403
Db	342	AsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLysGluArgMetCys	361
QY	1404	AGGCCTTCTTCAAGAAATGCGACCTCCCTGGAATATTGCAATGGATCATCTGCATCATGC	1463
Db	362	ArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGlySerSerAlaSerCys	381
QY	1464	CCAGAAAACCATATGTTACAGCTGGGCATPCCTGTGGACTGAATCAATGATCTGTATA	1523
Db	382	ProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGlnTyrIleCysIle	401
QY	1524	GATGGAGTTGTATGATGGGGATAAACAATGTACAGACACATTTGSCAAAGAAGTAGAG	1583
Db	402	AspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGlyLysGluValGlu	421
QY	1584	TTTGGCCCTTCAGAATGTTATCTCACCTTAAATTTCAAAGACTGATGTATCTGGAACCTGT	1643
Db	422	PheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspValSerGlyAsnCys	441
QY	1644	GGTATAAGTGATTTCAGGATACACAGATGTGAAGCTGACAATCTGCAGTGGGAAAAATTA	1703
Db	442	GlyLysSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGlnCysGlyLysLeu	461
QY	1704	ATATGTAATATGATAGTAAATTTTATTACAAATTTCCAAAGACCACTATTATTATGCC	1763
Db	462	IleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAlaThrIleIleTyrAla	481

```
QY 1764 AACATAAGTGGACATCTCTGCTGATTCGTGGAATTTGCCAGTGATCATGACAGACGCAA 1823
Db 482 AsnIleSerGlyHisLeuCyIleAlaValGluPheAlaSerAspHisAlaAspSerGln 501
QY 1824 AAGATGTGGATAAAGATGAACCTTCTGTGGTTCAAAATAAGGTTTCAGGAATCAAGA 1883
Db 502 LysMetIrpIleuysAspGlyThrSerCyseGlySerAsnLysValCySA-gAsnGlnArg 521
QY 1884 TGTGTGAGTCTTCATACCTTGGGTTATGATTTACTACTGACAAATGCAATGATAGAGGT 1943
Db 522 CysValSerSerSerTyLeuGlyTyAspCyseThrThrAspLysCysAsnAspArgGly 541
QY 1944 GTATGCAATTAACAAAAGCTGTCTACTGTAGTGCTTCATATTTACCTCCAGATGCTCA 2003
Db 542 ValCysAsnAsnLysHisCysHisCysSerAlaSerTyLeuProAspCysSer 561
QY 2004 GTTCAATCAGATCTATCGCTGGTGGGAGTATGACAGTGGCAATTTCCACCTGTAGCT 2063
Db 562 ValGlnSerAspLeuIrpProGlyGlySerIleAspSerGlyAsnPheProIvalAla 581
QY 2064 ATACCAAGCCAGACTCCCTGAAAGCGCTACATTCAGAACATTTTACCAATTCACAAACCAATG 2123
Db 582 IleProAlaArgLeuProGluArgArgTyIleGluAsnIleTyHisSerLysProMet 601
QY 2124 AGATGGCAATTTTCTTATTCATTCCTCTCTTATTTATTTATTTCTGTGTACTGCTATA 2183
Db 602 ArgTrpProPhePheLeuPheIleProPhePheIlePheCysValLeuIleAlaIle 621
QY 2184 ATGGTGAAGCTTAATTCACAAAGAAAATGGAGAACTGAGACTATTCACGGATGAG 2243
Db 622 MetValLysValAsnPheGlnArgLysLysTrpArgThrGluAspTySerSerAspGlu 641
QY 2244 CAACCTGAAAGTGAAGTGAACCTTAAAGGG 2273
Db 642 GlnProGluSerGluSerGluProLysGly 651

RESULT 6
US-08-765-243-6
; Sequence 6, Application US/08765243
; Patent No. 5935578
; GENERAL INFORMATION:
; APPLICANT: ALVES, KENNETH
; APPLICANT: GUPTA, SUNIL K.
; APPLICANT: HOLLI, GREGORY F.
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MARY A. APOLLINA
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,243
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: APOLLINA, MARY A
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19244Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-765-243-6

Alignment Scores:
Pred. No.: 7,98e-222 Length: 735
Score: 2440.00 Matches: 439
Percent Similarity: 75.71% Conservative: 119
Best Local Similarity: 59.57% Mismatches: 167
Query Match: 50.88% Indels: 12
DB: 2 Gaps: 7

US-10-054-683-18 (1-2640) x US-08-765-243-6 (1-735)
QY 78 GTCTGTTTCTGCTCAGCGGG-----CTCGGGGGCTG-----CGATGGAC 119
Db 4 IleLeuLeuLeuLeuSerGlyLeuSerGluLeuGlyGlyLeuSerGlnSerGlnThrGlu 23
QY 120 AGTAATTTTGTATGTTTACCTGCTGCAATTTACAGTTCCGAGAGAAATACCGTCAATAA 179
Db 24 GlyThrArgGluLysLeuHisValGlnValThrValProGluLysIleArgSerValThr 43
QY 180 AAGGAAGGAATTGAATCGCAGGCATCTACAAATTTGTAATTCAGAGGAAACCATATACT 239
Db 44 SerAsnGlyTyLeuThrGlnValThrTyAsnLeuLysIleGluGlyLysThrTyThr 63
QY 240 GTGAATTTTAAATGCAAAAAAACTTTTACCCCAATAATTTTAGAGTTTACAGTTTAGTGGC 299
Db 64 LeuAspLeuMetGlnLysProPheLeuProAsnPheArgValTySerTyAspAsn 83
QY 300 ACAGGAATTTAAGAACCATCTGACCAAGATTTTCAGAAATTTCCGCACCTACCAAGGTAT 359
Db 84 AlaGlyIleMetArgSerLeuGluGlnLysPheGlnAsnIleCysTyPheGlnGlyTy 103
QY 360 ATTGAAGGTTATCCAAAATCTGTGGTGATGTTAGCACATGTACTGTGACTCAGGGCGTA 419
Db 104 IleGluGlyTyProAsnSerMetValIleValSerThrCysThrGlyLeuArgGlyPhe 123
QY 420 CTACAGTTTGAAATGTTAGTTATGGAATAGAACCCCTGGAGTCTTCAGTTGGCTTTGAA 479
Db 124 LeuGlnPheGlyAsnValSerTyGlyIleGluProLeuGluSerSerGlyPheGlu 143
QY 480 CATGTAATTTTACCAAGTAAACATAGAAAGCAGATGTTTCCTTATATATATAGAAAGAT 539
Db 144 HisValIleTyLeuGlnValGluProGluLysGlyGlyAlaLeuLeuTyAlaGluLysAsp 163
QY 540 ATTGAATCAAGAGATCTGTCTCTTAAATTAATTAACAAGCGCAGACGACAGCAAGATTTGCA 599
Db 164 IleAspLeuArgAspSerGlnTyLysIleArgSerIleLysProGlnArgIleValSer 183
QY 600 AAGTATATAGAAATGCAATGTTATAGTTGAAAAACAATTTGATATATCATATGGGCTCGAT 659
Db 184 HisTyLeuGluIleHisIleValValGluLysGlnMetPheGluHisIleGlyAlaAsp 203
QY 660 ACAACTGTTGCTGCTCAAAAGTTTTCAGTTGATGGATTGACGATGCTATTTTGT 719
Db 204 ThrAlaIleValThrGlnLysIlePheGlnLeuIleGlyLeuAlaAsnAlaIlePheAla 223
QY 720 TCATTTAATTAATTAATTTCTGTTCTTCTTCTGAGCTTTGGATAGATAAAATAAAAT 779
Db 224 ProPheAsnLeuThrValIleLeuSerSerLeuGluPheTrpMetAspGluAsnLysIle 243
QY 780 GCAACCACTGGAGAAGCTAATAGTTTATACACATTTTAAAGATGGAAGAAACATCTTAT 839
Db 244 LeuThrThrGlyAspAlaAsnLysLeuLeuTyArgPheLeuLysTyLysGlnSerTy 263
QY 840 CTTGTTTACGCTCATGATGCGCATTTTACTTGTGTTTACAGAGAAAGTCAATAT 899
Db 264 LeuValLeuArgProHisAspMetAlaPheLeuLeuValTyArgAsnThrThrAspTy 283
QY 900 GTTGGTGAACCTTTCAAGGGAAGATGTGTGATGCAAACTATTCAGGAGGTGTTGTTCTG 959
Db 284 ValGlyAlaThrTyLeuGlnGlyLysMetCysAspLysAsnTyAlaGlyValAlaLeu 303
```

```
QY 960 CACCCAGAACCAATAAGCTCGGAATCACTTGCAGTATTTTACCTCAATTATTGACCTT 1019
Db 961 SerGlyAsn---LysGluArgAlaGluPro-----IleProValArgProTyrIleAla 677
QY 1020 AGTATGGGATCACTATGATGACATTAACAATAATGCGAGTCTCAGGAGCTGTCGATT 1079
Db 324 SerMetGlyLeuAlaTyrAspAspValAsnLysCysGlnCysGlyValProValCysVal 343
QY 1080 ATCAATCCCAAGACAAATTCATTTTCAGTGTGTGAAGATCTTTAGTAATCTCAGCTTCGAA 1139
Db 344 MetAsnProGluAlaProHisSerSerGlyValArgAlaPheSerAsnCysSerMetGlu 363
QY 1140 GACTTTGCACATTTATTTCAAGCAGAAAGTCCAGTGTCTTCAATACAGCTTCGCTTA 1199
Db 364 AspPheSerLysPheIleThrSerGlnSerSerHisCysLeuGlnAsnGlnProThrLeu 383
QY 1200 GATCCTTTTTCACAGCAGCAGTGTGTGTAATGCAAGCTGGAAGCAGGAGGAG 1259
Db 384 GlnProSerTyrLys---MetAlaValCysGlyAsnGlyValGluGluAspGluIle 402
QY 1260 TGTGACTGTGGGACTGAACAGGATTTGCCCTTATTGGAGAAACATGCTGTGATATTGCC 1319
Db 403 CysAspCysGly---LysLysGlyCysAlaGluMetProProCysCysAsnProAsp 421
QY 1320 ACATGTAGATTTAAAGCCGTTCAAACTGTGTGTAAGGACCATGCTCGGAAAACTGCTA 1379
Db 422 ThrCysLysLeuSerAspGlySerGluCysSerSerGlyIleCysCysAsnSerCysLys 441
QY 1380 TTTATGTCAAGAAAGATGTTGTAGCCCTTCCTTTGAAGAAATCGACCTCCCTGATAT 1439
Db 442 LeuLysArgLysGlyGluValCysArgLeuAlaGlnAspGluCysAspValThrGluTyr 461
QY 1440 TGCAATGGATCATCTGCATCATGCCAGAAACACATATGTCAGACTGGGCATCCGTGT 1499
Db 462 CysAsnGlyThrSerGluValCys---GluAspPheValGlnAsnGlyHisProCys 480
QY 1500 GGACTGAATCAATGGATCTGTATAGATGGAGTTTGTATGATGGGGATAAACAATGTACA 1559
Db 481 AspAsnArgLysTrpIleCysIleAsnGlyThrCysGlnSerGlyGluGlnCysGln 500
QY 1560 GACACATTTGGCAAGAGTAGATTTGGCCCTTCAGATGTTATCTACCTTAATTC 1619
Db 501 AspLeuPheGlyIleAspAlaGlyPheGlySerSerGluCysPheTrpGluLeuAsnSer 520
QY 1620 AAGACTGATGATCTGTGAAACTGTGTATAGTAAAGTATTTCAGATACACACAGTGTGAAGCT 1679
Db 521 LysSerAspIleSerGlySerCysGlyIleSerAlaGlyTyrLysGluCysProPro 540
QY 1680 GACAATCTGCAGTCGCGAAATTAATATGTAATATGATAGTAAATTTTATTACAAAT 1739
Db 541 AsnAspArgMetCysGlyLysIleIleCysLysTyrGlnSerGluAsnIleLeuLysLeu 560
QY 1740 CCAAGAGCCACTATTATTTATGCCAATAGTGGACATCTCTGCATTGCTGTGGAAATTT 1799
Db 561 ArgSerAlaThrValIleTyrAlaAsnIleSerGlyHisValCysValSerLeuGluTyr 580
QY 1800 GCCAGTGATCATGCAGACCCAAAGATGCTGGATAAAGATGAACTTCTGTGGTTCA 1859
Db 581 ProGlnGlyHisAsnGluSerGlnLysMetTrpValArgAspGlyThrValCysGlySer 600
QY 1860 AATAAGTTTTCAGGAAATCAAGATGTGTGAGTCTTTCATACATTTGGTGTATGTTGACT 1919
Db 601 AsnLysValCysGlnAsnGlnLysCysValAlaAspThrPheLeuGlyTyrAspCysAsn 620
QY 1920 ACTGACAAATGCAATGATAGAGTGTATGCAATAACAAAGACATGCTCACTGTAGTCT 1979
Db 621 LeuGluLysCysAsnHisHisGlyValCysAsnAsnLysLysAsnCysHisCysAspPro 640
QY 1980 TCATATTTTACCTCCAGATTCCTCAGTTCAATCAGATCTATGGCTGTGTGGAGTATTGAC 2039
Db 641 ThrTyrLeuProProAspCysLysArgMetLysAspSerTyrProGlyGlySerIleAsp 660
```

```
QY 2040 AGTGGCAATTTTCCACCTGTAGCTATATACAGCCAGACTCCTCGAAAGCCGCTACATTGAG 2099
Db 661 SerGlyAsn---LysGluArgAlaGluPro-----IleProValArgProTyrIleAla 677
QY 2100 AACATTTTACCATTCCAAACCAATGAGATGCGCAATTTTCTTATTCTTCTTCTTTTATT 2159
Db 678 SerArgTyrArgSerLysSerProArgTrpProPhePheLeuIleLeuProPheTyrVal 697
QY 2160 ATTTTCTGTGTACTGATTAATAATGGTGAAGTTAATTTCCAAAGGAAAAAATGGAGA 2219
Db 698 ValIleLeuValLeuIleGlyMetLeuValLysValTyrSerGlnArgMetLysTyrArg 717
QY 2220 ACTGAGGACTATTCAAGCAGATGAGCAACCTGAAAGTGAGAGTGAACTTAA 2270
Db 718 MetAspAspPheSerSerGluGluGlnPheGluSerGluSerGluSerLys 734

RESULT 7
PCT-US95-07295-6
; Sequence 6, Application PC/TUS9507295
; GENERAL INFORMATION:
; APPLICANT: ALVES, KENNETH
; APPLICANT: GUPTA, SUNIL K.
; APPLICANT: HOLLIS, GREGORY F.
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MARY A. APOLLINA
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07295
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: APOLLINA, MARY A
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19244Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-07295-6

Alignment Scores:
Pred. No.: 7,98e-222 Length: 735
Score: 2440.00 Matches: 439
Percent Similarity: 75.71% Conservative: 119
Best Local Similarity: 59.57% Mismatches: 167
Query Match: 50.88% Indels: 12
DB: 5 Gaps: 7

US-10-054-683-18 (1-2640) x PCT-US95-07295-6 (1-735)
QY 78 GTCTTGTCTTGTCTACGCGG-----CTCGCGGGCTG-----CGATGGAC 119
Db 4 IleLeuLeuLeuLeuSerGlyLeuSerGluLeuGlyGlyLeuSerGlnSerGlnThrGlu 23
QY 120 AGTAATTTTCATAGTTTACCTGTGCAATTTACAGTTCCGAGAAATACCGTCAATAATA 179
Db 24 GlyThrArgGluLysLeuHisValGlnValThrValProGluLysIleArgSerValThr 43
```



; APPLICANT: ALVES, KENNETH  
 ; APPLICANT: GUPTA, SUNIL K.  
 ; APPLICANT: HOLLIS, GREGORY F.  
 ; TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MARY A. APOLLINA  
 ; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE  
 ; CITY: RAHWAY  
 ; STATE: NJ  
 ; COUNTRY: USA  
 ; ZIP: 07065  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent-In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/264,101  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: APOLLINA, MARY A.  
 ; REGISTRATION NUMBER: 34,087  
 ; REFERENCE/DOCKET NUMBER: 19244  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (908)594-3462  
 ; TELEFAX: (908)594-4720  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 457 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-264-101-4

## Alignment Scores:

Pred. No.: 3,36e-134 Length: 457  
 Score: 1512.00 Matches: 265  
 Percent Similarity: 73.70% Conservative: 74  
 Best Local Similarity: 57.61% Mismatches: 115  
 Query Match: 31.53% Indels: 6  
 DB: 1 Gaps: 5

US-10-054-683-18 (1-2640) x US-08-264-101-4 (1-457)

QY 891 TCAATATGTTGGTCAACCTTTCAAGGGAAGATGTTGATCGCAAACTATGCGAGGTT 950  
 DB 3 SerAspTyrValGlyAlaThrTyrGlnGlyLysMetCysAspLysAsnTyrAlaGlyGly 22  
 QY 951 GTTGTCTCGACCCCAAGCAACATAAGTCTGGAATCACTTCAGTATTTTGTAGCTCAATTA 1010  
 DB 23 ValAlaLeuHisProLysAlaValThrLeuGluSerLeuAlaIleLeuValGlnLeu 42  
 QY 1011 TTGAGCCTTAGTATGGGATCACTTATGATGACATTAAACAAATGCCAGTCTCAGGAGCT 1070  
 DB 43 LeuSerLeuSerMetGlyLeuAlaTyrAspValAsnLysCysGlnCysGlyValPro 62  
 QY 1071 GTCTGCATTATGATCAGAAAGCAATTCATTGAGTGGTGTGAAGATCTTTTAGTAATGTC 1130  
 DB 63 ValCysValMetAsnProGluAlaProHisSerSerGlyValArgAlaPheSerAsnCys 82  
 QY 1131 AGCTTCGAAGACTTGCACATTTTATTCAAGCAGAAAGTCCAGTGTCTTCAACATCAG 1190  
 DB 83 SerMetGluAspPheSerLysPheIleThrSerGlnSerSerHisCysLeuGlnAsnGln 102  
 QY 1191 CCTCGCTTAGATCCTTTTTCAACACAGCAAGCTGTGTGTATGCAAAAGCTGGAAGCA 1250  
 DB 103 ProThrLeuGlnProSerTyrLys---MetAlaValCysGlyAsnGlyGluValGluGlu 121  
 QY 1251 GGAAGAGAGTGTGACTGTGGACTGAACACAGGATTTGTCCTTATTTGGAGAAACATGCTGT 1310  
 DB 1251 GGAAGAGAGTGTGACTGTGGACTGAACACAGGATTTGTCCTTATTTGGAGAAACATGCTGT 1310

## RESULT 9

US-08-765-243-4  
 ; Sequence 4, Application US/08765243  
 ; Patent No. 5935578  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALVES, KENNETH  
 ; APPLICANT: GUPTA, SUNIL K.

Db 122 AspGluIleCysAspCysGly---LysLysGlyCysAlaGluMetProProProCysCys 140  
 QY 1311 GATATTGCCACATGATGATTAAAGCGCGTTCAAACTGTCTGAAGACCATGCTGCGAA 1370  
 Db 141 AsnProAspThrCysLysLeuSerAspGlySerGluCysSerSerGlyLysCysCysAsn 160  
 QY 1371 AACTGCTCTATTATGTCAAAGAAAGAAATGTGTAGGCTTCTTCTTGAAGAAATGCGACCTC 1430  
 Db 161 SerCysLysLeuLysArgLysGlyGluValCysArgLeuAlaGlnAspGluCysAspVal 180  
 QY 1431 CCTGAATATTGCAATGATCATCTGCATCATGCCAGAAACCACCACTATGTTGCACATGGG 1490  
 Db 181 ThrGluTyrCysAsnGlyThrSerGluValCys---GluAspPhePheValGlnAsnGly 199  
 QY 1491 CATCCGTGTGGACTGAATCAATGATCTGTATAGATGGAGTTCATGATGAGTGGGATAAA 1550  
 Db 200 HisProCysAspAsnArgLysTrpIleCysIleAsnGlyThrCysGlnSerGlyGluGln 219  
 QY 1551 CAATGTACAGACACATTGTCGCAAGAGTAGAGTTTGGCCCTTCAGAAATGTTATTCTCAC 1610  
 Db 220 GlnCysGlnAspLeuPheGlyIleAspAlaGlyPheGlySerSerGluCysPheTrpGlu 239  
 QY 1611 CTTAATTCGAAGACTGATGATCTGGAACCTGTGGTATAGTATTGATTCAGGATACACACAG 1670  
 Db 240 LeuAsnSerLysSerAspIleSerGlySerCysGlyIleSerAlaGlyGlyTyrLysGlu 259  
 QY 1671 TGTGAAGCTGACAACTGTCAGTGCAGAAATTAATATGTAATATGATGAGTAAATTTTA 1730  
 Db 260 CysProAsnAspArgMetCysGlyLysIleIleCysLysTyrGlnSerGluAsnIle 279  
 QY 1731 TTACAAATTCGAAGACCACTATTATTATGCCAAATAGTGGACACTCTCTGATGCT 1790  
 Db 280 LeuLysLeuArgSerAlaThrValIleTyrAlaAsnIleSerGlyIleValCysValSer 299  
 QY 1791 GTGGAATTTGCCAGTGCATGACAGCAGCAAGATGTTGGTAAAGATGGAACCTTCT 1850  
 Db 300 LeuGluTyrProGlnGlyHisAsnGluSerGlnLysMetTrpValArgAspGlyThrVal 319  
 QY 1851 TGTGGTTTCAATTAAGGTTTGCAGAAATCAAGATGTTGAGTCTTCTCATCTCTGGGTAT 1910  
 Db 320 CysGlySerAsnLysValCysGlnAsnGlnLysCysValAlaAspThrPheLeuGlyTyr 339  
 QY 1911 GATTGTACTACTGACAAATGCAATGATAGAGGTGTATGCAATTAACAAAGACACTGCTAC 1970  
 Db 340 AspCysAsnLeuGluLysCysAsnHisGlyValCysAsnAsnLysLysAsnCysHis 359  
 QY 1971 TGATGCTTCATATTTACCTCCAGATGCTCAGTTCATCAATCAGATCTATGCCCTGGTGG 2030  
 Db 360 CysAspProThrTyrLeuProProAspCysLysArgMetLysAspSerTyrProGlyGly 379  
 QY 2031 AGTATTGACAGTGGCAATTTCCACCTGTAGCTATACCCAGCAGACTCCCTGAAAGCGC 2090  
 Db 380 SerIleAspSerGlyAsn---LysGluArgAlaGluPro-----IleProValArgPro 396  
 QY 2091 TACATTGAGAACATTTACCAATCCAAACCAATCAGATGCGCACTTTTCTTATTCTCTCT 2150  
 Db 397 TyrIleAlaSerArgTyrArgSerLysSerProArgTrpProPhePheLeuIlePro 416  
 QY 2151 TTCTTTATTTTCTGTGTACTGATGCTATATGTTGAAGTAAATTTTCCAAAGGAAA 2210  
 Db 417 PheTyrValValIleLeuValIleGlyMetLeuValLysValTyrSerGlnArgMet 436  
 QY 2211 AAATGAGAACTGAGGACTATTCAAGCGATGAGCAACTCAAGATGAGAGTGAACCTAAA 2270  
 Db 437 LysTrpArgMetAspAspPheSerSerGluGlnPheGluSerGluSerLys 456







```

CORRESPONDENCE ADDRESS:
ADDRESSEE: MARY A. APOLLINA
STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07295
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: APOLLINA, MARY A
REGISTRATION NUMBER: 34,087
REFERENCE/DOCKET NUMBER: 19244Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)594-3462
TELEFAX: (908)594-4720
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 457 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-07295-4

Alignment Scores:
Pred. No.: 3.36e-134
Score: 1512.00
Percent Similarity: 73.70%
Best Local Similarity: 57.61%
Query Match: 3.53%
DB: 5

US-10-054-683-18 (1-2640) x PCT-US95-07295-4 (1-457)

```









Db	210	LeuGluLeuValValValValAAsnHisAAspPhePheIleTyrSerGlnSerAAsnIleSer	229
Qy	666	GTTCCTCGCTCAAAAAGTTTTCCAGTTGATGGATGAGCAATGCTATTGTTGTTTCATTT	725
Db	230	LysValGlnGluAspValPheLeuValValAAsnIleValAAsnSerMetTyrGlnGlnLeu	249
Qy	726	AATATTACAATTATTCTGCTTCATTGGAGCTTTGGATAGATGAAATAAAATTCGAACC	785
Db	250	GlyThrTyrIleIleLeuIleGlyIleGluIleTyrAsnGlnGlyAAsnValPheProMet	269
Qy	786	ACTGGAGAAGCTAATGAGTTATTACACACATTTTTAAGATGGAAAAACATCTTATCTTGT	845
Db	270	Thr---SerIleGluGlnValLeuAAsnAspPheSerGlnTrrIysGlnIleSerLeuSer	288
Qy	846	TTAGCTCCTCATGATGGCATTTTACTTTGTTTACAGAAAAAGTCAAATATTGTTGGT	905
Db	289	GlnLeuGlnHisAspAlaAlaHisMetPheIleGlyAAsnSerIleLeuSerIleLeuGly	308
Qy	906	GCAACCTTTCAAGGGAAGATGTGTGATGCA-----AACTATGCA	944
Db	309	LeuAlaTyrValAlaGlyIleCysArgProProlIleAspCysGlyValAAsnPheGln	328
Qy	945	GGAGGTGTTGTTCTGCAACCCAGACACATTAAGCTTGGATCATCTTCGAGTTATTTTAGCT	1004
Db	329	GlyAsp-----ThrTrrSerLeu-----PheAlaAsnThrValAla	340
Qy	1005	CAATTATTGAGCCTTAGTATGGGATCAGCTTATGATGACATTAAACAATGCCAGTGTCTCA	1064
Db	341	HisGluLeuGlyHisThrLeuGlyMetGlnHisAspGlu---GluPheCysPheCysGly	359
Qy	1065	GGAGCTGCTGCAATTATGAATCCAGAAGCAATTCATTTCAGTGGTGAAGATCTTTAGT	1124
Db	360	GluArgGlyCysIleMetAsnThrPheArgVal-----ProAlaGluLysPheThr	376
Qy	1125	AACCTGACGCTCGAAGACTTTGCACATTTTATTTCAAAGCAGAAGTCCCAAGTCTTTCAC	1184
Db	377	AsnCysSerTyrAlaAspPheMetLysThrThrLeuAAsnGlnGlySer---CysLeuHis	395
Qy	1185	AATCAGCCTCGCTTAGATCCTTTTTCAAACAGCAGCAGTGTGTGGTATGCAAGACTG	1244
Db	396	AsnProProArgLeuGlyGluIlePheMetLeuLysArg---CysGlyAAsnGlyValVal	414
Qy	1245	GAAGCAGGAGAGAGTGTGACTGTGGAGCTGAACAGGATGTGCCCTTATTGGAGAAACA	1304
Db	415	GluArgGluGlnGlnCysAspCysGlySerValGlnGlnCysGlu-----GlnAspAla	432
Qy	1305	TGCTGTGATATTGCCACATAGATTTAAAGCCGGTTCAAACCTGCTCGTAGGACCATTGC	1364
Db	433	CysCys---LeuLeuAAsnCysThrLeuAArgProGlyAlaAlaCysAlaPheGlyLeuCys	451
Qy	1365	TGGAAACACTGCTCTATTATGTCAAAAGAAGAATGTGTAGGCCCTTCCTTTGAAGATGC	1424
Db	452	CysIleAspCysLysPheMetProSerGlyGluLeuCysArgGlnGluValAAsnGlyCys	471
Qy	1425	GACCTCCCTGAATATTGCAATGATCATCTGCATCATGCCACAGAAACCACTATGTTTCAG	1484
Db	472	AspLeuProGlnTrrCysAAsnGlyThrSerHisGlnCysProGluAAspArgTyrValGln	491
Qy	1485	ACTGGGCATCCGTGTGGACTGAATCAATGATCTGTATAGATGGAGTTGTATGATGGGG	1544
Db	492	AspGlyIleProCysSerAspSerAlaTyr---CysTyrGlnLysArgCysAAsnAAsnHis	510
Qy	1545	GATAAACAATGTTACAGACACATTTGGCAAGAAGTAGAGTTTGGCCCTTCAGAAATGTTAT	1604
Db	511	AspGlnHisCysArgGluIlePheGlyLysAspAlaLysSerAlaSerGlnAAsnCysTyr	530
Qy	1605	TCTCACCTTAATTCAAAGACTGATGTATCTGGAAACTGTGGTAAGTAGTATTCAGGATAC	1664
Db	531	LysGluIleAAsnSerGlnGlyAAsnArgPheGlyHisCysGlyIleAAsnGlyThrThrTyr	550
Qy	1665	ACACGTGTGAAGCTGACAAATCTGCACTGCGGAAAAATTAATATGTAAATATGTAGGTAA	1724

```
Db 551 LeuLysCysHisIleSerAspValPheCysGlyArgValGlnCysGluAenValArgAsp 570
QY 1725 TTTTATTACAAATTCAGAGCCACTATTATTATGCCAACATAAGTGGACATCTCTGC 1784
Db 571 IleProLeuLeuGlnAspHisPheThrLeuGlnHisThrHisIleAenGlyValThrCys 590
QY 1785 ATTGCTGTGGAATTT-----GCCAGTGCATCATGCACAGACCCAAAGATG 1829
Db 591 TrpGlyIleAspTyrHisLeuArgMetAsnIleSerAspIleGlyGlu----- 606
QY 1830 TGGATAAAGACGAACTCTTGTTGTTCAAAATAGGTTTGCAGGAATCAAGATGGTG 1889
Db 607 ---ValLysAspGlyThrValCysGlyProGlyLysIleCysIleHisLysCysVal 625
QY 1890 AGTCTTCATPACTGGGTTATGATGTACTACTGACAAATGCAATGATAGAGGTATGC 1949
Db 626 SerLeuSerValLeuSerHisValCysLeuProGluThrCysAenMetLysGlyIleCys 645
QY 1950 AATAACAAAAGCACTGCTCACTGTAGTCTTATTTATTTACTCCAGATTGCTCAGTTCAA 2009
Db 646 AsnAsnLysHisHisCysHisCysGlyTyrGlyTrpSerProProTyrCysGlnHisArg 665
QY 2010 TCAGATCATGGCTGTGGAGTATTGACAGTGGCAATTTTCCACCTGTAGTATACCA 2069
Db 666 Gly-----TyrGlyGlySerIleAspSerGly-----Pro 675
QY 2070 GCCAGACTCCCTGAAAGGCGCTACATTGAGAACAATTTTACCATTCCAAACCAATCAGATGG 2129
Db 676 Ala-SerAlaLysArgGlyValPheLeuProLeuIleValIlePro----- 690
QY 2130 CCATTTTCTATTCACTCTCTCTCTTTTATTATTCTGTG 2169
Db 691 -----SerLeuSerValLeuThrPheLeuPheThrVal 701
```

## RESULT 15

```
US-09-786-256C-15
; Sequence 15, Application US/09786256C
; Patent No. 6680189
; GENERAL INFORMATION:
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: HIKIKI, Yuichi
; APPLICANT: NISHIMURA, Atsushi
; TITLE OF INVENTION: No. 6680189el Protein and DNA Thereof
; FILE REFERENCE: 2544 USOP
; CURRENT APPLICATION NUMBER: US/09/786,256C
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: PCT/JP99/04766
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: JP 10-250115
; PRIOR FILING DATE: 1998-09-03
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 15
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(775)
; OTHER INFORMATION: An isolated ADAM family protein
US-09-786-256C-15
```

```
Alignment Scores:
Pred. No.: 6 4e-71 Length: 775
Score: 843.50 Matches: 210
Percent Similarity: 45.59% Conservative: 121
Best Local Similarity: 28.93% Mismatches: 300
Query Match: 17.59% Indels: 95
DB: 4 Gaps: 23
```

US-10-054-683-18 (1-2640) x US-09-786-256C-15 (1-775)

QY 156 CCGGAGAAAATACGGTCAATAATAAGGAAGGAATTGAATCGCAGGCATCTACAAAATT 215

```
Db 47 ProGluGln-----GlnGluGlnPheGluThrGluLeuLysTyrLysMet 61
QY 216 GTAATTAGAGGAAACCATATATCTGTAATTTAATGCAA---AAAAATCTTTTACCCCAT 272
Db 62 ThrIleAsnGlyLysIleAlaValLeuTyrLeuLysLysAsnLysLeuLeuAlaPro 81
QY 273 AATTAGAGTTTACAGTTATAGTCGACACAGAAATTATGAACACCTTGACCAAGATTTT 332
Db 82 GlyTyrThrGluThrTyrTyrAsnSerThrGlyLysGluIleThrThrSerProGlnIle 101
QY 333 CAGAAATTTCTGCCACTACCAAGGGTATATGAAGTTATCCAAAATCTGGTGGTATGTT 392
Db 102 MetAspAspCysTyrTyrGlnGlyHisIleLeuAsnGluLysValSerAspAlaSerIle 121
QY 393 AGCACATGTACTCGACTCAGGGCGCTACTACAGTTTGAAATTTAGTTATGGAATAGAA 452
Db 122 SerThrCysArgGlyLeuArgGlyTyrPheSerGlnGlyAspGlnArgTyrPheIleGlu 141
QY 453 CCCTCGAG-----TCTTCAGTTGGCTTTGAACATGTAAT-----TACCAAGTAAAA 500
Db 142 ProLeuSerProIleHisArgAspGlyGlnGluHisAlaLeuPheLysTyrAsnProAsp 161
QY 501 CATAAGAACAGCATGTTCTTATATATGAAGGATATTGAATCAAGAGATCTCTCC 560
Db 162 GluLysAsnTyrAspSerThrCysGlyMetAspGlyValLeuTrpAlaHisAspLeuGln 181
QY 561 -----TTTAAATTCAAAAGCGCAGACCCACACCAA 590
Db 182 GlnAsnIleAlaLeuProAlaThrLysLeuValLysLeuLysAspArgLys---ValGln 200
QY 591 GATTTCGCAAGTATATAGAAATGTCATGTTATAGTTGAAAACAAATTTGTAATAATCATATG 650
Db 201 GluHisGluLysTyrIleGluTyrTyrLeuValLeuAspAsnGlyGluPheLysArgTyr 220
QY 651 GGCTCTGATACAACTGTTGCTCGCTCAAAAAGTTTTCAGTTGATGATGATGACAAATGCT 710
Db 221 AsnGluAsnGlnAspGluIleArgLysArgValPheGluMetAlaAsnTyrValAsnMet 240
QY 711 ATTTTGTGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 770
Db 241 LeuTyrLysLysLeuAsnThrHisValAlaLeuValGlyMetGluIleThrPheAspLys 260
QY 771 AATAAAATTCGAACACCTGAGAGCTAATGATGATTAATCACAACATTTTAAAGATGAAA 830
Db 261 AspLysIleLysIleThrProAsnAlaSerPheThrLeuGluAsnPheSerLysTrpArg 280
QY 831 ACATCTTATCTGTTTACGTCTCT---CATGATGGCATTGTTTACTGTTTACAGAA 887
Db 281 GlySerValLeuSerArgArgLysArgHisAspIleAlaGlnLeuIleThrAlaThrGlu 300
QY 888 ---AAGTCAAAATTAATGTTGTCGAACCTTTCAAGGGAAGATGTTGATGCAAACTATGCA 944
Db 301 LeuAlaGlyThrThrValGlyLeuAlaPheMetSerThrMetCys---SerProTyrSer 319
QY 945 GGAGGTGTTGTTCTGCAACCCCAACCAATAAGTCTGGAATCACTTGCAGTTATTTAGCT 1004
Db 320 ValGlyValVal---GlnAspHisSerAspAsnLeuLeuArgValAlaGlyThrMetAla 338
QY 1005 CAATTATTGAGCCTTAGTATGGGATCACTTATGATGACATTAAACAAATCCAGTGTCTCA 1064
Db 339 HisGluMetGlyHisAsnPheGlyMetPheHisAspAspTyrSer---CysLysCysPro 357
QY 1065 GGAGCTGCTGTCATTATGAATCCAGAAGCAATTCATTCTAGTGGTGGTGAAGATCTTAGT 1124
Db 358 SerThrIleCysValMetAsp---LysAlaLeuSerPheTyrIleProThrAspPheSer 376
QY 1125 AACTGCAGCTTCGAAGACTTTGCACATTTTATTTCAAAAGCAGAAAGTCCAGTGTCTTAC 1184
Db 377 SerCysSerArgLeuSerTyrAspLysPhePheGluAspLysLeuSerAsnCysLeuPhe 396
QY 1185 AATCAACCTCGCTTAGATCCTTTTTTCAAAACAGCAAGCAGTGTGTGTAATGCAAAAGCTG 1244
```

Search completed: January 10, 2005, 22:09:27  
Job time : 122 secs

```
Db 397 AsnAlaPro---LeuProThrAspIleSerThrProIleCysGlyAsnGlnLeuVal 415
QY 1245 GAAGCAGGAGGAGGAGTGTGACTGTGGGACTCAACAGGATTGTGCCCTTATTGGAGAAACA 1304
Db 416 GluMetGlyGluAspCysAspCysGlyThrSerGluGluCysThrAsnIle----- 432
QY 1305 TGCTGTGATATGCGCATGATAGATTTAAAGCCGGTTCAAACTGTGCTGAAGGACCATGC 1364
Db 433 CysCysAspAlaIleThrCysIleLysAlaThrPheGlnCysAlaLeuGlyGluCys 452
QY 1365 TGGCAAACTGTCTATTATTGTCAAAGAAGATGTAGGCTTCCTTTGAAGATGC 1424
Db 453 CysGluLysCysGlnPheLysLysAlaGlyMetValCysArgProAlaLysAspGluCys 472
QY 1425 GACTCCCTCAATATTGCAATCATCTGCATCTGCCAGAAACCACTATGTTTCAG 1484
Db 473 AspLeuProGluMetCysAsnGlyLysSerGlyAsnCysProAspAspArgPheGlnVal 492
QY 1485 ACTGGGCATCCGTGTGGACTGAATCAATGGATCTGTATAGATGGAGTTTGTATGAGTGGG 1544
Db 493 AsnGlyPheProCysHisHisGlyLysGlyHisCysLeuMetGlyThrCysProThrLeu 512
QY 1545 GATAAACAATGTACAGACACATTGGCAAGAAGTAGAGTTTGGCCCTTCAGATGTTAT 1604
Db 513 GlnGlnCysThrGluLeuTrpGlyProGlyThrGluValAlaAspLysSerCysIle 532
QY 1605 TCTCACCTTAATTCAAAGACTGTATCTGGAACCTGTGGTATAGTATTCAGGATAC 1664
Db 533 AsnArg---AsnGluGlyGlySerLysTyrglyTyrcysargargvalaspThrLeu 551
QY 1665 ACACAGTGTGAAGCTGACAATCTGCAGTGGCGGAAATAATATATGTA--- 1712
Db 552 IleProCysLysAlaAsnAspThrMetCysGlyLysLeuPheCysGlnGlyGlySerAsp 571
QY 1713 -----TATGTAGTAATTTTATTACAAATTCAGAGCCACTATTATTATGCC 1763
Db 572 AsnLeuProTrpLysGlyArgIleValThrPheLeuThrCysLysThr----- 587
QY 1764 AACATAAGTGGACATCTCTGCATTGCTGTGGAATTTGCCAGTGTATCATGCAGACAGCAA 1823
Db 588 -----PheAspProGluAspThrSerGlnGlu 596
QY 1824 AGATGTGGATAAAGATGAACCTTCTGTGGTTCAATAGGTTTGCAGGAATCAAAGA 1883
Db 597 IleGlyMetValAlaAsnGlyThrLysCysGlyAspAsnLysValCysIleAsnAlaGlu 616
QY 1884 TGTCTG-----AGTCTTCTCATCTTGGTTATCATTTGACTACTGACAAATGCAATGAT 1937
Db 617 CysValAspIleGluLysAlaIleThrLysSerThrAsnCysSerSer---LysCysLysGly 635
QY 1938 AGAGGTGTATGCAATAACAAAGCACTGTCACTGTAGTGTCTCATATTTACCTCCAGAT 1997
Db 636 HisAlaValCysAspHisGluLeuGlnCysGlnCysGluGluGlyTrpIleProAsp 655
QY 1998 TGCTCAGTTCAATCAGATCTATGGCCTGGTGGGAGTATGACAGTGGCAATTTTCCACCT 2057
Db 656 CysAspAspSerSer----- 660
QY 2058 GTAGCTATACCGCCAGACTCCCTGAAAGGGCTACATTGGAACATTTACCATTCCAA 2117
Db 661 -----ValValPheHisPheSer 666
QY 2118 CCAATGAGATGGCCATTTTCTTATTATTCATCTCTTTCTTTATTATTATTCTGTGACTGATT 2177
Db 667 IleValValGlyValLeuPhe-----ProMetAlaValIlePhe---ValValVal 682
QY 2178 GCTATAATGGTGAAGTTAATTTCCAAAGGAAAAATGGAGAACTGAGGACTATTCAAGC 2237
Db 683 AlaMetValIleArgHisGlnSerSerArgGluLysGlnLys-----Lys 697
QY 2238 GATCAGCAACTGAACT 2255
Db 698 AspGlnArgProLeuSer 703
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 10, 2005, 21:35:41 ; Search time 328.5 Seconds  
(without alignments)  
5795.484 Million cell updates/sec

Title: US-10-054-683-18

Perfect score: 4796

Sequence: 1 catctgcgacttcacactgc.....actgttacattcacaacaaa 2640

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1603904 seqs, 360571292 residues

Total number of hits satisfying chosen parameters: 3207808

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USFTO.spool\_p/US10054683/runat\_10012005\_174512\_15537/app\_query.fasta\_1.2823  
-DB=Published Applications\_AA -QFMT=fascan -SURFIX=n2p.rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0  
-NCPU=6 -ICPU=3 -NMAB=LARGEQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Published Applications AA:
- 1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/prodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/prodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/prodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/prodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/prodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/prodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/prodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/prodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 20: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

\* Query  
Result

No.	Score	Match	Length	DB	ID	Description
1	3984	83.1	734	14	US-10-054-683-19	Sequence 19, Appl
2	3984	83.1	734	14	US-10-205-823-10	Sequence 10, Appl
3	3984	83.1	734	17	US-10-473-603-19	Sequence 19, Appl
4	3969.5	82.8	735	14	US-10-205-823-8	Sequence 8, Appl
5	3900	81.3	753	14	US-10-205-823-6	Sequence 6, Appl
6	1530	31.9	820	10	US-09-961-656-2	Sequence 2, Appl
7	1524.5	31.8	787	14	US-10-227-884-90	Sequence 90, Appl
8	1524.5	31.8	787	14	US-10-230-163-90	Sequence 90, Appl
9	1524.5	31.8	787	14	US-10-230-338-90	Sequence 90, Appl
10	1524.5	31.8	787	14	US-10-218-631-90	Sequence 90, Appl
11	1524.5	31.8	787	14	US-10-230-414-90	Sequence 90, Appl
12	1524.5	31.8	787	14	US-10-232-224-90	Sequence 90, Appl
13	1524.5	31.8	787	14	US-10-216-159A-90	Sequence 90, Appl
14	1524.5	31.8	787	14	US-10-218-849-90	Sequence 90, Appl
15	1524.5	31.8	787	14	US-10-227-873-90	Sequence 90, Appl
16	1524.5	31.8	787	14	US-10-227-883-90	Sequence 90, Appl
17	1524.5	31.8	787	14	US-10-219-076-90	Sequence 90, Appl
18	1524.5	31.8	787	14	US-10-230-434-90	Sequence 90, Appl
19	1524.5	31.8	787	14	US-10-219-003-90	Sequence 90, Appl
20	1524.5	31.8	787	14	US-10-219-075-90	Sequence 90, Appl
21	1524.5	31.8	787	14	US-10-219-464-90	Sequence 90, Appl
22	1524.5	31.8	787	14	US-10-219-466-90	Sequence 90, Appl
23	1524.5	31.8	787	14	US-10-219-479-90	Sequence 90, Appl
24	1524.5	31.8	787	14	US-10-219-481-90	Sequence 90, Appl
25	1524.5	31.8	787	14	US-10-230-260-90	Sequence 90, Appl
26	1524.5	31.8	787	14	US-10-232-231-90	Sequence 90, Appl
27	1524.5	31.8	787	14	US-10-232-233-90	Sequence 90, Appl
28	1524.5	31.8	787	14	US-10-216-165-90	Sequence 90, Appl
29	1524.5	31.8	787	14	US-10-218-956-90	Sequence 90, Appl
30	1524.5	31.8	787	14	US-10-219-468-90	Sequence 90, Appl
31	1524.5	31.8	787	14	US-10-219-478-90	Sequence 90, Appl
32	1524.5	31.8	787	14	US-10-219-536-90	Sequence 90, Appl
33	1524.5	31.8	787	14	US-10-233-205-90	Sequence 90, Appl
34	1524.5	31.8	787	14	US-10-219-072-90	Sequence 90, Appl
35	1524.5	31.8	787	14	US-10-219-474-90	Sequence 90, Appl
36	1524.5	31.8	787	14	US-10-219-524-90	Sequence 90, Appl
37	1524.5	31.8	787	14	US-10-219-528-90	Sequence 90, Appl
38	1524.5	31.8	787	14	US-10-219-528-90	Sequence 90, Appl
39	1524.5	31.8	787	14	US-10-227-880-90	Sequence 90, Appl
40	1524.5	31.8	787	14	US-10-227-881-90	Sequence 90, Appl
41	1524.5	31.8	787	14	US-10-232-223-90	Sequence 90, Appl
42	1524.5	31.8	787	14	US-10-230-436-90	Sequence 90, Appl
43	1524.5	31.8	787	14	US-10-232-225-90	Sequence 90, Appl
44	1524.5	31.8	787	14	US-10-232-227-90	Sequence 90, Appl
45	1524.5	31.8	787	14	US-10-232-227-90	Sequence 90, Appl

ALIGNMENTS

RESULT 1

US-10-054-683-19  
; Sequence 19, Application US/10054683  
; Publication No. US20030044813A1  
; GENERAL INFORMATION:  
; APPLICANT: Old, Lloyd J.  
; APPLICANT: Scanlan, Matthew J.  
; APPLICANT: Chen Yao-Tseng  
; TITLE OF INVENTION: Cancer-Testis Antigens  
; FILE REFERENCE: L0461/7125 (JRV)  
; CURRENT APPLICATION NUMBER: US/10/054, 683  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: US 60/280, 718  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 60/285, 154  
; PRIOR FILING DATE: 2001-04-20  
; PRIOR APPLICATION NUMBER: US 60/327, 432  
; PRIOR FILING DATE: 2001-10-05  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 734



```
QY 2052 CCACCTGTAGCTATACAGCCAGACTCCCTGAAAGGGCTACATTGAGAACATTTTACCAT 2111
Db 661 ProProValAlaileProAlaargLeuProGluArgArgTyrileGluAenileTyrHis 680
QY 2112 TCAAAACCAATGAGTGGCCATTTTCTTATTCATTCCTTTCTTTATTTCTGTGTA 2171
Db 681 SerLysProMetArgTrpProPheLeuPhePhePhePhePhePhePhePhePheCysVal 700
QY 2172 CTGATTGCTATAATGGTGAAGTTAAATTTCCAAAGGAAAAAATCGAGAACTGAGGACTAT 2231
Db 701 LeuileAlaileMetVallyseValAenPheGlnArglyslsyrpArgThrGluAspTyr 720
QY 2232 TCAAGCGATGAGCAACTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2273
Db 721 SerSerAspGluGlnProGluSerGluSerGluSerGluProlysgly 734

RESULT 2
US-10-205-823-10
; Sequence 10, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsay, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-10
Alignment Scores:
Pred. No.: 0 Length: 734
Score: 3984.00 Matches: 734
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.07% Indels: 0
DB: 14 Gaps: 0

US-10-054-683-18 (1-2640) x US-10-205-823-10 (1-734)
QY 72 ATGTGGTCTGTGTCTCTGCTCAGCGGCTCGCGGGCTCGGATGAGTAAATTTTGCAT 131
Db 1 MetTrpValLeuPheLeuLeuSerGlyLeuGlyLeuArgMetAspSerAsnPheAsp 20
QY 132 AGTTTACCTGTGCAATACAGTTCGGAGAAAATACGGTCAATATAAAGGAGGAATT 191
Db 21 SerLeuProValGlnileThrValProGluLysileArgSerileileLysGluGlyile 40
```

```
QY 192 GAATCCAGGCATCTCTACAAAAATTGTAATTGAAGGAAACCATATACTGTGAATTTAATG 251
Db 41 GluSerGlnAlaSerTyrLysileValileGluGlyLysProTyrThrValAsnLeuMet 60
QY 252 CAAAAAATCTTTTACCCCATATTTTACAGTTTACAGTTTATAGTGGCAGCAGNAATTTATG 311
Db 61 GlnLysAsnPheLeuProHisAsnPheArgValTyrSerTyrSerGlyThrGlyileMet 80
QY 312 AAACCACTTGACCAAGATTTTCAGAAATTTCTGCCACTTACCAAGGGTATATTGAAGGTAT 371
Db 81 LysProLeuAspGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrileGluGlyTyr 100
QY 372 CCAAAATCTGTGTGTATGTTAGCATGTACTGTGACTCAGCGGCTGACTACAGTTTGAA 431
Db 101 ProLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGlu 120
QY 432 AATGTTAGTTATGGAATAGNACCCCTGGAGTCTTTCAGTTGGCTTTCGAACTGTAATTTAC 491
Db 121 AsnValSerTyrGlyileGluProLeuGluSerSerValGlyPheGluHisValileTyr 140
QY 492 CAAATAAACATAAGAAAGCAGATGTTTCTTATATAATGAGAAGATATTGAATCAAGA 551
Db 141 GlnValylHisLysLysAlaAspValSerLeuTyrAsnGlnLysAspLileGluSerArg 160
QY 552 GATCTCTCTTTAAATTTACAAAGCGCAGAGCCACAGCAAGATTTTGCAAAGTATATAGAA 611
Db 161 AspLeuSerPheLysLeuGlnSerAlaGluProGlnAspPheAlaLysTyrileGlu 180
QY 612 ATGCATGTTATAGTTGAAAAACAATTTGTATATCATATGCGGCTGTGATACAACTGTTGTC 671
Db 181 MetHisValileValGluLysGlnLeuTyrAsnHisMetGlySerAspThrThrValVal 200
QY 672 GCTCAAAAGTGTTCAGTTGATTGAGTGCAGCAATGCTATTTTGTTCATTTAATATT 731
Db 201 AlaGlnLysValPheGlnLeuileGlyLeuThrAsnAlailePheValSerPheAsnile 220
QY 732 ACAATATTCTGTCTTTCATTGGAGCTTTGGATAGATGAAATATAAATTTGCAACACCTGGA 791
Db 221 ThrileileLeuSerSerLeuGluLeuTrpIleAspGluAenLysileAlaThrThrGly 240
QY 792 GAAGCTAATGAGTTATACACATTTTAAAGATGGAACATCTTATCTTGTGTTTACGT 851
Db 241 GluAlaAsnGlnLeuLeuHisThrPheLeuArgTrpLysThrSerTyrLeuValLeuArg 260
QY 852 CCTCATGATGTGTCATTTTACTTGTTCAGAGAAAGTCAAAATATGTTGGTGCACACC 911
Db 261 ProHisAspValAlaPheLeuLeuValTyrArgGlnLysSerAsnTyrValGlyAlaThr 280
QY 912 TTTCAAGGGAAGATGTGTGATGCAAACTATGCAAGGAGGTGTGTTCTGCAACCCAGAAC 971
Db 281 PheGlnGlyLysMetCysAspAlaAsnTyrAlaGlyValValLeuHisProArgThr 300
QY 972 ATAAGTCTGGAATCATTGCGAGTTATTTAGCTCAATTTATGAGCTTAGTATGGGGATC 1031
Db 301 IleSerLeuGluSerLeuAlaValileLeuAlaGlnLeuLeuSerLeuSerMetGlyile 320
QY 1032 ACTTATGATGACATTAACAAATCCAGTGTCTCAGGAGCTGTCTGCAATTATGAAATCCAGAA 1091
Db 321 ThrTyrAspAspLileAsnLysCysGlnCysSerGlyAlaValCysileMetAsnProGlu 340
QY 1092 GCAATTCATTTCAGTGTGTGAAGATCTTTTAGTAACCTGCAAGCTTCGAGCTTTGCAACAT 1151
Db 341 AlaileHisPheSerGlyVallyllePheSerAsnCysSerPheGluAspPheAlaHis 360
QY 1152 TTTATTTCAAGCAGAGTCCCGAGTGTCTTACAACTCAGCTCGCTTAGTATCTTTTTC 1211
Db 361 PheileSerlysglnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhePhe 380
QY 1212 AAACACACAGCAGTGTGTGTATGTAATGAAAGCTCGAAGCAGAGGAGTGTGCTGTGGG 1271
Db 381 LysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCysGly 400
```

```
QY 1272 ACTGAACAGGATGTGGCCCTTATTGGAGAAACATGCTGTGATATTGCCACATGTAGATT 1331
Db ThrGluGlnAspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPhe 420
QY 1332 AAAGCCGTTCAAACTGTCTGAAGGACATGCTGGGAAACTGTCTATTATTATGTCAAAA 1391
Db LysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLys 440
QY 1392 GAAAGAATGTGTAGGCCCTTCTTGAAGAATGCCACCTCCCTGATATTGGCAATGGATCA 1451
Db GluArgMetCysArgProSerPheGluGluCysAspLeuProGluLysCysAsnGlySer 460
QY 1452 TCTCATCATGCCAGAAACCACTATGTTCCAGACTGGGCATCCGTGTGGACTGAATCAA 1511
Db SerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGln 480
QY 1512 TGGATCTGTATAGATGGATTTGTATGATGGGGATAAACAATGTACAGACACATTTGGC 1571
Db TrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGly 500
QY 1572 AAAGAAGTAGATTTGGCCCTTCAAGATGTTATTCTCACCTTAATTCAAAGACTGATGA 1631
Db LysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspVal 520
QY 1632 TCTGAAACTGTGTATAAGTATGATTCAGGATACACACAGTGTGAAGCTGCAATCTGCAG 1691
Db SerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGln 540
QY 1692 TCGGAAATTAATATGTAATATGATAGTAAATTTTATTATCAAAATCCAGAGCCACT 1751
Db CysGlyLysLeuIleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAlaThr 560
QY 1752 ATTATTTATGCCAACATAAGTGGACATCTCTGCATTTGCTGTGGAATTTGCCAGTGCAT 1811
Db IleIleTyrAlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHis 580
QY 1812 GCAGACAGCCAAAGATGTGGATAAAGATGGAATCTTCTGTGTTCAAATTAAGTTTGC 1871
Db AlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCys 600
QY 1872 AGGAATCAAGATGTGTGATTTCTTCACTCTGCTGATTTGATTTGATCTACTGCAAAATGC 1931
Db ArgAsnGlnArgCysValSerSerTyrLeuGlyTyrAspCysThrThrAspLysCys 620
QY 1932 AATGATAGAGTGTATGCAATAACAAAGACCTGTCTGCTGTGTCTTCAATTTTACCT 1991
Db AsnAspArgGlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeuPro 640
QY 1992 CCAGATTGCTCAGTTCAATCAGATCTATGGCTGTGGAGTATGTACAGTGGCAATTT 2051
Db ProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsnPhe 660
QY 2052 CCACCTGTAGCTATACCAGCCAGACTCCCTGAAAGCGCTACATTTGAGACATTTACCAT 2111
Db ProProValAlaIleProAlaArgLeuProGluArgTyrIleGluAsnIleTyrHis 680
QY 2112 TCAAAACCAATGATGGCCATTTTCTTATTCATTCCTTCTTATTATTATTTCTGTGTA 2171
Db SerLysProMetArgTrpProPhePheLeuPheIleProPheIleIlePheCysVal 700
QY 2172 CTGATTCCTATATGTGGAAGTAAATTTCAAAGGAAATAATGGAGAACTGAGGACTAT 2231
Db LeuIleAlaIleMetValLysValAsnPheGlnArgLysLysTrpArgThrGluAspTyr 720
QY 2232 TCAAGCATGACGAACTGAAAGTGAAGTGAACCTAAAGGG 2273
Db SerSerAspGluGlnProGluSerGluSerGluProLysGly 734
```

RESULT 3

US-10-473-603-19

; Sequence 19, Application US/10473603

; Publication No. US20040235066A1

; GENERAL INFORMATION:

```
; APPLICANT: OLD, Lloyd
; APPLICANT: SCANLAN, Matthew
; APPLICANT: CHEN, Yao-Tseng
; TITLE OF INVENTION: CANCER-TESTIS ANTIGENS
; FILE REFERENCE: L0461.70155US00
; CURRENT APPLICATION NUMBER: US/10/473.603
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: PCT/US02/09808
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 10/054.683
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/280.718
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/285.154
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/327.432
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-473-603-19
```

```
Alignment Scores:
Pred. No.: 0 Length: 734
Score: 3984.00 Matches: 734
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.07% Indels: 0
DB: 17 Gaps: 0
```

US-10-054-683-18 (1-2640) x US-10-473-603-19 (1-734)

```
QY 72 ATGTGGTCTTGTCTTCTGCTCAGCGGCTCGCGGCTCGGATGGACAGTAATTTTGAT 131
Db 1 MetTrpValLeuPheLeuLeuSerGlyLeuGlyLeuArgMetAspSerAsnPheAsp 20
QY 132 AGTTTACCTGTGCAAAATACAGTTCGGAGAAATACGTCATATAAAGGAAGAAAT 191
Db 21 SerLeuProValGlnIleThrValProGluLysIleArgSerIleIleLysGluGly 40
QY 192 GAATCCGAGCATCTTACAAAATTTAAATTTGAAGGAAACCATATACTGTGAATTTAATG 251
Db 41 GluSerGlnAlaSerTyrLysIleValIleGluGlyLysProTyrThrValAsnLeuMet 60
QY 252 CAAAAAATCTTTTACCCCATAAATTTTAGAGTTTACAGTTTATAGTGGCACAGGAATATG 311
Db 61 GlnLysAsnPheLeuProHisAsnPheArgValTyrSerTyrSerGlyThrGlyIleMet 80
QY 312 AAACCACTTGACCAAGATTTTCAGAAATTTCTGCCACTACCAAGGTATATTGAAGTTAT 371
Db 81 LysProLeuAspGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGlyTyr 100
QY 372 CAAAAATCTGTGTGATGTTAGTACCATGTCTGAGCTCAGGGGCGCTACTACAGTTTGA 431
Db 101 ProLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGlu 120
QY 432 AATGTTAGTATTGGAATAGAACCCCTGGAGTCTTTCAGTTGGCTTTGAAACATGAAATTAC 491
Db 121 AsnValSerTyrGlyIleGluProLeuGluSerSerValGlyPheGluHisValIleTyr 140
QY 492 CAAGTAAAAACATAAGAAACAGATTTTCTCTATATAATGAGAAGGATATTGAAATCAAGA 551
Db 141 GlnValLysHisLysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSerArg 160
QY 552 GATCTCTCTCTTAAATTTACAAAGCGCAGAGCCACAGCAAGATTTTGCAAAGTATATAGAA 611
Db 161 AspLeuSerPheLysLeuGlnSerAlaGluProGlnGlnAspPheAlaLysTyrIleGlu 180
QY 612 ATGCATGTTATAGTTCAAAAAACAATTTGATATCATATCGGTCTGATCAACTGTTGTC 671
|||
```

181 MetHisValIleValGluLysGlnLeuTyrAsnHisMetGlySerAspThrThrValVal 200  
QY GCTCAAAAGATTTTCCAGTTGATTGGATTCACGAATCGTATTTTGTTCATTAAATTT 731  
Db AlaGlnLysValPheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsnIle 220  
QY ACAATTATTTCTGCTTCATTTGGAGCTTTGGATAGATCAAAATTAATTTGCAACCACTGGA 791  
Db ThrIleIleLeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThrThrGly 240  
QY GAAAGCTAATCAGTATTACACACATTTTAAAGTATGGAATCATCTTATCTTTGTTTACGT 851  
Db GluAlaAsnGlnLeuLeuHisThrPheLeuArgTrpLysThrSerTyrLeuValLeuArg 260  
QY CCTCATGATGCGCATTTTACTGTTTACAGAAAGTCAAAATTAATTTGTTGTCGAACC 911  
Db ProHisAspValAlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAlaThr 280  
QY TTTCAAGGGAAGATGTTGATGCANAATATGCAGGAGGTGTTGTTCTGCACCCAGNACC 971  
Db PheGlnGlyLysMetCysAspAlaAsnTyrAlaGlyValValLeuHisProArgThr 300  
QY ATAAGTCTGGAATCAGTTGAGTATTATTTAGCTCAATTTAGTACCTTAGTATGGGATC 1031  
Db IleSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGlyIle 320  
QY ACTTATGATCAGATTAAACAAATCCAGTCTGCAGGAGCTCTGCATTTATGAATCCAGAA 1091  
Db ThrTyrAspAspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnProGlu 340  
QY GCAATTCATTTAGTGTGTGAAGATCTTTAGTAACTGCAGAGCTTCGAGACTTCGAGACTTTGCACAT 1151  
Db AlaIleHisPheSerGlyValValLysIlePheSerAsnCysSerPheGluAspPheAlaHis 360  
QY TTTTATTTCAAGCAGAGTCCAGTGCTTCAATTCAGCTCGCTTGTAGATCTCTTTTTC 1211  
Db PheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhePhe 380  
QY AARACAGACAGTGTGTGTAATGCAAGCTGGAAGCAGAGAGGAGTGTGATGTGGG 1271  
Db LysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCysGly 400  
QY ACTGACAGAGATTGTCCTTTATTTGAGAAACATGCTGTGATATTGACCATGTAGATTT 1331  
Db ThrGluGlnAspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPhe 420  
QY AAAGCCGGTTCAAACTGCTGAGGACCATGCTGCGAAACTGCTCTATTATTATGTCAAAA 1391  
Db LysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLys 440  
QY GAAAGAAATGTGTAGGCCCTTCCTTTGAAGAATGCGACCTCCCTCAATATTGCAATGGATCA 1451  
Db GluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGlySer 460  
QY TCTGCATCATGCCAGAAAACCATATGTTTCAGACTGGGATCCGTTGAGTGAATCAAA 1511  
Db SerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGln 480  
QY TGGATCTGTATAGATGAGTTTCTGATGAGTGGGATAAACAATGTCACAGACATTTGGC 1571  
Db TrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGly 500  
QY AAAGAAGTAGAGTTTGGCCCTTCAGAAATGTTTATTCACCTTAATTCAAAGACTGTATGTA 1631  
Db LysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspVal 520  
QY TCTGGAACCTGTGTATAGTATTCAGGATACACAGTGTGAAGTGTGAATCTGCAG 1691  
Db SerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGln 540  
QY TCGGGAATAATATGTAATATGATGTAATTTTATACAAATTCGAAGGCCACT 1751  
Db CysGlyLysLeuIleCysTyrValGlyLysPheLeuLeuGlnIleProArgAlaThr 560

QY ATTATTTATGCCAACATAAGTGGACATCTCTGCATTGCTGTGGAATTTGCCAGTGATCAT 1811  
Db IleIleTyrAlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHis 580  
QY GCAGACAGCCAAAGATGTGATAAAGATGGAACTTCTTGTGTTCAAAATAGGTTTGC 1871  
Db AlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCys 600  
QY AGAAATCAAGATGTGTGAGTCTTTCATCTCTGGTTTATGATTGTACTACTGACAAATGC 1931  
Db ArgAsnGlnArgCysValSerSerSerTyrLeuGlyTyrAspCysThrThrAspLysCys 620  
QY AATGATAGAGGTGTATGCAATAACAAAAGACCTGTCACCTGTAGTGTCTTATTTACCT 1991  
Db AsnAspArgGlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeuPro 640  
QY CCAGATTGCTCAGTTCAATCAGATCTATGSCCTGTGGAGTATTGACAGTGGCAATTTT 2051  
Db ProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsnPhe 660  
QY CCACCTGTAGCTATACAGCCAGACCTCCCTGAAAGCGCTACATTTGAGAACATTTACCAT 2111  
Db ProProValAlaIleProAlaArgLeuProGluArgTyrIleGluAsnIleTyrHis 680  
QY TCCAAACCAATGAGATGGCCATTTTCTTATTCATTCCTCTTCTTATTTATTTCTGTGTA 2171  
Db SerLysProMetArgTrpProPhePheLeuPheIleProPheIleIlePheCysVal 700  
QY CTGATTGCTATATGTTGAAGTTAATTTCCAAAGGAAAATGGAGAACTGAGGACTAT 2231  
Db LeuIleAlaIleMetValLysValAsnPheGlnArgLysLysTrpArgThrGluAspTyr 720  
QY TCAAGCGATGAGCAACTGAAAGTGAGAGTGAAACCTAAAGGG 2273  
Db SerSerAspGluGlnProGluSerGluSerGluProLysGly 734

RESULT 4  
US-10-205-823-8  
; Sequence 8, Application US/10205823  
; Publication No. US20030108963A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Gorbacheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Wonsey, Angela M.  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Anderson, Dustin  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
; FILE REFERENCE: MRI-044  
; CURRENT APPLICATION NUMBER: US/10/205,823  
; CURRENT FILING DATE: 2002-07-25  
; PRIOR APPLICATION NUMBER: 60/307,982  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: 60/314,356  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/325,020  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: 60/341,746  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/362,158  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 455  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 735

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-8

Alignment Scores:
Pred. No.: 0
Score: 3969.50
Percent Similarity: 99.73%
Best Local Similarity: 99.73%
Query Match: 82.77%
DB: 14

Length: 735
Matches: 733
Conservative: 0
Mismatches: 1
Indels: 1
Gaps: 1

US-10-054-683-18 (1-2640) x US-10-205-823-8 (1-735)

QY 72 ATGTGG---GTCCTGTTCTCTCAGCGGGCTCGCGGGCTCGAGTGCAGCAAGTAATTTT 128
DB 1 MetTrpArgValLeuPheLeuLeuSerGlyLeuGlyGlyLeuArgMetAspSerAsnPhe 20
QY 129 GATAGTTTACCTGTCACAAATTTACAGTTCGCGAGAAATACGGTCAATAATAAGGAAGGA 188
DB 21 AspSerLeuProValGlnIleThrValProGluLysIleArgSerIleLeuLysGluGly 40
QY 189 ATTGAATCGCAGGATCTCTACAAATTTGTAATTTGAAGGAAACCATATACGTGAATTA 248
DB 41 IleGluSerGlnAlaSerTyrLysIleValIleGluGlyLysProTyrThrValAsnLeu 60
QY 249 ATGCAAAAAAATTTTACCCCAATTTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTT 308
DB 61 MetGlnLysAsnPheLeuProHisAsnPheArgValTyrSerTyrSerGlyThrGlyLe 80
QY 309 ATGAAACCACTGTACCAAGATTTTTCAGAAATTTTCCCACTACCAAGGATATTTGAAGGT 368
DB 81 MetLysProLeuAspGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGly 100
QY 369 TATCAAAATCTGCTGATGGTTAGCACATGTACTGACCTCAGCGGGCTACTACAGTTT 428
DB 101 TyrProLysSerValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPhe 120
QY 429 GAAATGTAGTTATGAATAAGAACCCCTGAGAGCTTCAGTTGGCTTTGAACATGTAATT 488
DB 121 GluAsnValSerTyrGlyIleGluProLeuGluSerSerValGlyPheGluHisValIle 140
QY 489 TACCAAGTAAACATAAGAACAGATGTTTCCCTTATATATATGAAGAGATATTTGAATCA 548
DB 141 TyrGlnValLysHisLysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSer 160
QY 549 AGAGATCTGCTTTAAATTTACAAAGCCGACAGCCACAGCAGATTTTGCAGATATATA 608
DB 161 ArgAspLeuSerPheLysLeuGlnSerValGluProGlnAspPheAlaLysTyrIle 180
QY 609 GAAATGCATGTTATAGTTGAAAAACAATTTGTAATCATATGAGGCTCTGATACAACTGTT 668
DB 181 GluMetHisValIleValGluLysGlnLeuTyrAsnHisMetGlySerAspThrVal 200
QY 669 GTCGCTCAAAAGTTTCCAGTTGATGGAATGCAAGATGCTATTTTGTGTTCAATTTAAT 728
DB 201 ValAlaGlnLysValPheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsn 220
QY 729 ATTCAATATTCTGCTCTTCAATTCAGCTTTGGATAGATGAAATTAATTTGCACCACT 788
DB 221 IleThrIleLeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThr 240
QY 789 GGAGAACTAATGAGTTTATACACACATTTTAAAGATGGAACAATCTTATCTTGTGTTTA 848
DB 241 GlyGluAlaAsnGluLeuLeuHisThrPheLeuArgTrpLysThrSerTyrLeuValLeu 260
QY 849 CGTCCTCATGATGGCATTTTACTGTTTACAGAAAGTCAAAATATGTTGGTGCA 908
DB 261 ArgProHisAspValAlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAla 280
QY 909 ACCTTTCAGAGGAGATGTGTGATGCAACATATTCAGGAGGTGTTGTTCTGCCCCCGAGA 968
DB 281 ThrPheGlnGlyLysMetCysAspAlaAsnTyrAlaGlyGlyValValLeuHisProArg 300
```

```

969 ACCATAAGTCTGGAATCAGTTGCAAGTTATTTAGCTCAATTTATTAGAGCTTAGTATGGG 1028
301 ThrIleSerLeuGluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGly 320
1029 ATCACTTATGATGACATTAAACAAATGCCAGTGTCTCAGAGCTGTCTCATTATGAATCCA 1088
321 IleThrTyrAspAspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnPro 340
1089 GAAGCAATTCATTTTCAGTGTGTGAAGATCTTTAGTAACTGCAGCTTCGAGACTTTGCA 1148
341 GluAlaIleHisPheSerGlyValIlePheSerAsnCysSerPheGluAspPheAla 360
1149 CATTTTATTTCAAAGCAGAAAGTCCCAGTGTCTTCACAATCAGCTCGCTTGTAGTCCCTTT 1208
361 HisPheIleSerLysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspPhe 380
1209 TTCAAACAGCAAGCAGTGTGTGTAATTCGAAAGCTGGAAGCAGAGAGAGTGTGACTGT 1268
381 PheLysGlnGlnAlaValCysGlyAsnAlaLysLeuGluAlaGlyGluCysAspCys 400
1269 GGGACTGAACAGGATTTGTGCCCTTATTGGAGAAACATGCTGTGTGATATTGCCACATGAGA 1328
401 GlyThrGluGlnAspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArg 420
1329 TTTAAAGCCGGTTCAAACCTGCTGAAGGACCATGCTGCCAAACTGCTCTATTATTGTCA 1388
421 PheLysAlaGlySerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSer 440
1389 AAAGAAGAAATGTGTAGGCTTCTCTTTGAAGAATGGACCTCCCTGTAATATTGCAATGGA 1448
441 LysGluArgMetCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGly 460
1449 TCATCTGCATCATGCCAGAAACCACTATGTTTCAGACTGGGCTCGCTGTGACTCAAT 1508
461 SerSerAlaSerCysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsn 480
1509 CAATGATCTGTATAGATGAGTTTGTATGAGTGGGATAAACAATGTACAGACACATTT 1568
481 GlnTrpIleCysIleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPhe 500
1569 GGCAAAAGATAGAGTTTGGCTTCCCTTCAAGATGTTTCTCACCTTAAATTCAAAGACTGAT 1628
501 GlyLysGluValGluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAsp 520
1629 GTATCTGGAACCTGCTGATTAAGTATAGGATACACAGTGTGAAGTGCACATCTG 1688
521 ValSerGlyAsnCysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeu 540
1689 CAGTGGCGAAAAATTAATATGTAATAATGATAGGTAAATTTTATTACAAATTTCCAAGAGCC 1748
541 GlnCysGlyLysLeuIleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAla 560
1749 ACTATTATTTATGCCAACATAAGTGACATCTCTGCAATGCTGTGGAATTTGCCAGTAT 1808
561 ThrIleIleTyrAlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAsp 580
1809 CATGCAGACAGCCAAAGATGTGGATAAAGATGGAACCTTCTGTGTTCAAAATAGGTT 1868
581 HisAlaAspSerGlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysVal 600
1869 TGCAAGAATCAAAGATGTGTGAGTCTTCTCATCTTGGGTTATGATTGCTACTCTGACAAA 1928
601 CysArgAsnGlnArgCysValSerSerTyrLeuGlyTyrAspCysThrThrAspLys 620
1929 TGCATATGATAGAGGTGTATGCAATAACAAAGACATGTCACGTGTAGTGTCTCATATTTA 1988
621 CysAsnAspArgGlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeu 640
1989 CCTTCCAGATTGCTCAGTTCAATCAGATCTATGCTGCTGGGAGTATTTCACAGTGCAT 2048
641 ProProAspCysSerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsn 660
```

```
QY 2049 TTTCCACCTGCTAGCTATACAGCCAGACTCCCTGAAAGGCGCTACATTGAGACATTTC 2108
Db 661 PheProValAlaIleProAlaAArgLeuProGluArgArgTyrIleGluAsnIleTyr 680
QY 2109 CATTCAAACCAATGAGATGCGCATTTTCTTATTCATTCTCTTCTTCTTATTTCTGT 2168
Db 681 HisSerLysProMetArgTrpProPhePheLeuPheIleProPhePheIleIlePheCys 700
QY 2169 GTACTGATTCCTATAATGGTGAAGTTAATTTCCAAAGGAAAAAATGGAGAACTGAGGAC 2228
Db 701 ValLeuIleAlaIleMetValLysValAsnPheGlnArgLysLysTrpAspThrGluAsp 720
QY 2229 TATTCAAGCATGAGCACTGAAAGTGAGAGTGACCTAAAGG 2273
Db 721 TyrSerSerAspGluGlnProGluSerGluSerGluProLysGly 735

RESULT 5
US-10-205-823-6
; Sequence 6, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 753
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-6

Alignment Scores:
Pred. No.: 0 Length: 753
Score: 3900.00 Matches: 720
Percent Similarity: 97.44% Conservative: 4
Best Local Similarity: 96.90% Mismatches: 13
Query Match: 81.32% Indels: 6
DB: 14 Gaps: 1

US-10-054-683-18 (1-2640) x US-10-205-823-6 (1-753)
QY 45 TCCGGCTGGGACCGAGACTTCAAGCCATGTGGGTCTTGTTCGTCAAGCGGCTCGG 104
Db 17 AlaAlaTrpGluLysGlyAlaHis-----IleProGlnArgLeuGly 30
QY 105 GGGCTCGGATGGACAGTAAATTTGATAGTTTACCTGTGCAATTAAGTTCCGGAGAAA 164
Db 31 LysLeuTyrThrAspGlnAspPheAspSerLeuProAlaGlnIleThrValProGluLys 50
```

```
QY 165 ATACGGTCAATATATAAGGAGGAATTGAATCCAGGCATCTTACAAAATTTGTAATTGAA 224
Db 51 IleArgSerIleIleLysGluGlyIleGluSerGlnAlaSerTyrLysIleValIleGlu 70
QY 225 GGGAAACCATATATCTGTAATTAATGCAAAAAAATTTTACCCTAATATTTAGAGTT 284
Db 71 GlyLysProTyrThrValAsnLeuMetGlnLysAsnPheLeuProHisAsnPheArgVal 90
QY 285 TACAGTTATAGTGGCAGCAGGAATTTATGAAACCACTTGCACCAAGATTTTCAGAATTTCTGC 344
Db 91 TyrSerTyrSerGlyThrGlyIleMetLysProLeuAspGlnAspPheGlnAsnPheCys 110
QY 345 CACTACCAAGGATATATTGAGTTTATCCTGATCTGCTGATGTTTACACATGTTACT 404
Db 111 HisTyrGlnGlyTyrIleGluGlyTyrProLysSerValValMetValSerThrCysThr 130
QY 405 GGACTCAGGGCGCTACTACAGTTTGAATTTAGTTATGTAATAGAACCCCTGGAGTCT 464
Db 131 GlyLeuArgGlyValLeuGlnPheGluAsnValSerTyrGlyIleGluProLeuGluSer 150
QY 465 TCAGTTGGCTTTGCAACATGTAATTTACCAAGTAAACATAGAAGAGCAGATGTTTCTCTTA 524
Db 151 SerValGlyPheGluHisValIleTyrGlnValLysHisLysLysAlaAspValSerLeu 170
QY 525 TATAATGAGAGGATATTTGAATCAAGAGATCTGCTCTTTAAATTTACAAAGCGCAGACCA 584
Db 171 TyrAsnGluLysAspIleGluSerArgAspLeuSerPheLysLeuGlnSerValGluPro 190
QY 585 CAGCAGATTTTGCAGAGTATATAGAAATGTCATGTTATAGTTTGAACCAATTTGTAAT 644
Db 191 GlnGlnAspPheAlaLysTyrIleGluMetHisValIleValGluLysGlnLeuTyrAsn 210
QY 645 CATATGGGCTCATGATCAACTGTTGCGCTCAAAAAGTTTTCAGATTTGATTTGACG 704
Db 211 HisMetGlySerAspThrThrValValAlaGlnLysValPheGlnLeuIleGlyLeuThr 230
QY 705 AATGCTATTTTGTTCATTATATATCAATTAATTAATTTCTCTTCATTGGAGCTTTGATA 764
Db 231 AsnAlaIlePheValSerPheAsnIleThrIleLeuSerSerLeuGluLeuTrpIle 250
QY 765 GATGAAATTAATTTGCACCACTGGAGAGCTTAATGAGTTATACACATTTTAAAGA 824
Db 251 AspGluAsnLysIleAlaThrThrGlyGluAlaAsnGluLeuHisThrPheLeuArg 270
QY 825 TCGAAAAACATCTTATCTTGTGTTTACGTCCTCATGATGTCGCTTTTCTTTTACAGA 884
Db 271 TrpLysThrSerTyrLeuValLeuArgProHisAspValAlaPheLeuLeuValTyrArg 290
QY 885 GAAAAGTCAAATTTATGTTGGTCAACCTTTTCAAGGGAAGATGTTGATGCAAACTATGCA 944
Db 291 GluLysSerAsnTyrValGlyAlaThrPheGlnGlyLysMetCysAspAlaAsnTyrAla 310
QY 945 GGAGGTGTTGTTCTGCAACCCAGAACCATTAAGTCTGGAATCACTTGCAGTATTTTACCT 1004
Db 311 GlyGlyValValLeuHisProArgThrIleSerLeuGluSerLeuAlaValIleLeuAla 330
QY 1005 CAATTATTGAGCTTAGTATGGGGATCACATTATGATGACATTAACAAATGCCAGTGTCTCA 1064
Db 331 GlnLeuLeuSerLeuSerMetGlyIleThrTyrAspAspIleAsnLysCysGlnCysSer 350
QY 1065 GGAGCTGTCTGCAATTTATGAATCCAGAGCAATTCATTTCAGTGGTGTGAAGATCTTTAGT 1124
Db 351 GlyAlaValCysIleMetAsnProGluAlaIleHisPheSerGlyValLysIlePheSer 370
QY 1125 AACTGAGCTCGAGACTTTTGACATTTTATTTCAAGAGAGAGTCCAGTGTCTTCAC 1184
Db 371 AsnCysSerPheGluAspPheAlaHisPheIleSerLysGlnLysSerGlnCysLeuHis 390
QY 1185 AATCAGCTCGCTTAGATCTCTTTTCAACAGCAAGCAGTGTGTGTTAATGCAAAAGCTG 1244
Db 391 AsnGlnProArgLeuAspProPhePheLysGlnAlaValCysGlyAsnAlaLysLeu 410
```



```
QY 1245 GAACAGGAGAGAGTGTGACTGTGGGACTGACACAGGATTGTGCCCTTTATTGGAGAAACA 1304
Db 411 GluAlaGlyGluGluCysAspCysGlyThrGluGlnAspCysAlaLeuIleGlyGluThr 430
QY 1305 TGCTGTGATATGCCACATGTAGATTAAAGCCGGTTCAAACCTGTGCTGAAGGACCATGC 1364
Db 431 CysCysAspIleAlaThrCysArgPheLysAlaGlySerAsnCysAlaGluGlyProCys 450
QY 1365 TGGCAAACTCTCTATTATGTGCAAAAGAAAGATGTGTAGCCCTCTCTTGAAGAAATGC 1424
Db 451 CysGluAsnCysLeuPheMetSerLysGluArgMetCysArgProSerPheGluGluCys 470
QY 1425 GACCTCCCTGAATATTGCAATGGATCATCTGCATCATCCAGAGAAACCATATGTTTCAG 1484
Db 471 AspLeuProGluTyrCysAsnGlySerSerAlaSerCysProGluAsnHisTyrValGln 490
QY 1485 ACTGGGCATCCGTGTGGACTGAATCAATGGATCTGTATAGATGGAAGTTGTATGAGTGG 1544
Db 491 ThrGlyHisProCysGlyLeuAsnGlnTrpIleCysIleAspGlyValCysMetSerGly 510
QY 1545 GATAAACAATGTACAGACACATTTGGCAAGAGATGTGTGGCCCTTCAGAATGTTAT 1604
Db 511 AsplysGlnCysThrAspThrPheGlyLysGluValGluPheGlyProSerGluCysTyr 530
QY 1605 TCTCACCTTAATCAAGACTGATGTATCTGGAACCTGTGTATAGTATGATTCAGGATAC 1664
Db 531 SerHisLeuAsnSerLysThrAspValSerGlyAsnCysGlyIleSerAspSerGlyTyr 550
QY 1665 ACACAGTGTGAAGCTGACAATCTGCAGTCCGGAATTAATATGTAATATGTTAGTTAAA 1724
Db 551 ThrGlnCysGluAlaAspAsnLeuGlnCysGlyLysLeuIleCysLysTyrValGlyLys 570
QY 1725 TTTTATTACAAATTCAGAGCCACTATTATTATGTCACATCAATAGTGGACATCTCTGC 1784
Db 571 PheLeuLeuGlnIleProArgAlaThrIleIleTyrAlaAsnIleSerGlyHisLeuCys 590
QY 1785 ATTGCTGTGAAATTTGCCAGTGATCATGCACAGCCCAAAAGATGTGTGATAAAGATGGA 1844
Db 591 IleAlaValGluPheAlaSerAspHisAlaAspSerGlnLysMetTrpIleLysAspGly 610
QY 1845 ACTCTGTGTGTTCAAAATAGGTTGTCAGGATCAAGATGTGTGAGTCTTCTACTATTG 1904
Db 611 ThrSerCysGlySerAsnLysValCysArgAsnGlnArgCysValSerSerTyrLeu 630
QY 1905 GGTATGATTGTACTGACAAATGCAATGATAGAGTGTATGCAATAACAAAGACAC 1964
Db 631 GlyTyrAspCysThrAspLysCysAsnAspArgGlyValCysAsnAsnLysLysHis 650
QY 1965 TGTCACTGTAGTCTTCATATTTTACCTCCAGATTGCTCAGTTCAATCAGATCTATGGCCT 2024
Db 651 CysHisCysSerAlaSerTyrLeuProProAspCysSerValGlnSerAspLeuTrpPro 670
QY 2025 GGTGGAGATTGACAGTGGCAATTTCCACTGTAGCTATACAGCCAGACTCCCTGAA 2084
Db 671 GlyGlySerIleAspSerGlyAsnPheProProValAlaIleProAlaArgLeuProGlu 690
QY 2085 AGGCGCTACATTGAGAAACATTACCATTTCCAAACCAATGAGATGCCATTTTCTTATTC 2144
Db 691 ArgArgTyrIleGluAsnIleTyrHisSerLysProMetArgTrpProPheLeuPhe 710
QY 2145 ATTCTCTTCTTATTATTTCTGTACTGATGCTATATGTTGGTAAAGTTAATTTCCAA 2204
Db 711 IleProPhePheIleIlePheCysValLeuIleAlaIleMetValLysValAsnPheGln 730
QY 2205 AGGAAAAATGGAGAACTGAGGACTATTCAAGCGATGAGCAACCTGAAAGTGAGAGTGAA 2264
Db 731 ArgLysLysTrpArgThrGluAspTyrSerSerAspGluGlnProGluSerGluSerGlu 750
QY 2265 CCTAAAGGG 2273
Db 751 ProLysGly 753
```

RESULT 6

```
US-09-961-656-2
; Sequence 2, Application US/09961656
; Publication No. US20030017572A1
; GENERAL INFORMATION:
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 56294, 56629, NOVEL HUMAN
; TITLE OF INVENTION: METALLOPROTEASES AND USES THEREOF
; FILE REFERENCE: 10448-092002
; CURRENT APPLICATION NUMBER: US/09/961,656
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/235,035
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 820
; TYPE: PRP
; ORGANISM: Homo sapiens
; US-09-961-656-2

Alignment Scores:
Pred. No.: 1,27e-126 Length: 820
Score: 1530.00 Matches: 318
Percent Similarity: 57.84% Conservative: 132
Best Local Similarity: 40.87% Mismatches: 266
Query Match: 31.90% Indels: 62
DB: 10 Gaps: 20

US-10-054-683-18 (1-2640) x US-09-961-656-2 (1-820)
QY 6 CGCATTTCCAACTGCCCTGTACACCACTGCCCTTTATTCGGCTGGAGCCAGGACTT 65
Db 22 ArgThrArgGlyCys-----TrpGlnProArgSer 31
QY 66 CAAGCCATGTGGTCTTGTGTTT---CTGCTCAGCGGCTCGCGGCTCGCGATGGAC--- 119
Db 32 ArgThrMetPheArgLeuTrpLeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArg 51
QY 120 AGTAATTTGATGTTTACCTGTGCAATTTACAGTTCCGAGAAATACCGTCAATATA 179
Db 52 ProGlyPheGlnAsnSerLeuLeuGlnIleValIleProGluLysIleGlnThrAsnThr 71
QY 180 AAGCAAGCAATTCAAA-----TCGACGCGCATCTTACAAAATTTGTAATTGAAGGAAA 230
Db 72 AsnAspSerSerGluIleGluTyrGluGlnIleSerTyrIleIleProIleAspGluLys 91
QY 231 CCATATCTGTGAATTAATGCAAAAACCTTTTACCCCATATTTTAGAGTTTACAGT 290
Db 92 LeuTyrThrValHisLeuLysGlnArgTyrPheLeuAlaAspAsnPheMetIleTyrLeu 111
QY 291 TATAGTGGCAGCAAGATTATGAACCACTTGACCAAGATTTTTCAGATTTTCCCATAC 350
Db 112 TyrAsn---GlnGlySerMetAsnThrTyrSerSerAspIleGlnThrGlnCysTyrTyr 130
QY 351 CAAGGGTATATCAAGGTTATCCAAATCTGTGGTGTATGCTAGCATGTCACATGTTCTG 410
Db 131 GlnGlyAsnIleGluGlyTyrProAspSerMetValThrLeuSerThrCysSerGlyLeu 150
QY 411 AGGGCGGTACTACAGTTTGAATGTTAGTTATGGAATAGAACCCCTCGGAGTCTTCAGT 470
Db 151 ArgGlyIleLeuGlnPheGluAsnValSerTyrGlyIleGluProLeuGluSerAlaVal 170
QY 471 GGCTTTGAACATGTAATTTACCAAGTAATAACATAGAAGCAGATGTTCTCTATATAAT 530
Db 171 GluPheGlnHisValLeuTyrLysLeuLysAsnGluAspAsnAspIleAlaIlePheIle 190
QY 531 GAGAAGGATATTGAATCAACA-----GATCTGTCTCTTAAATTTACAAAGCGCAGAG 581
Db 191 AspArgSerLeuLysGluGlnProMetAspAsnIlePheIleSerGluLysSerGlu 210
QY 582 CCACAGCAAGAT-----TTTGCAAGATATATAGAAATGCATGTTATAGTTGAAAAACA 635
```





;  
; PRIOR APPLICATION NUMBER: 60/069873  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082804  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/084441  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/085323  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/086392  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089905  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090691  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/095302  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/095318  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/095916  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: 60/096146  
; PRIOR FILING DATE: 1998-08-11  
; PRIOR APPLICATION NUMBER: 60/096791  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 60/097986  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: 60/098544  
; PRIOR FILING DATE: 1998-08-31  
; PRIOR APPLICATION NUMBER: 60/099596  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099598  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099803  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099811  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099812  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099816  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/100038  
; PRIOR FILING DATE: 1998-09-11  
; PRIOR APPLICATION NUMBER: 60/100385  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/100390  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/100627  
;  
;  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100848  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/100919  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/101477  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101738  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101741  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101786  
; PRIOR FILING DATE: 1998-09-25  
; PRIOR APPLICATION NUMBER: 60/101916  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101922  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/106178  
; PRIOR FILING DATE: 1998-10-28  
; PRIOR APPLICATION NUMBER: 60/106248  
; PRIOR FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: 60/106464  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 60/106905  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: 60/108787  
; PRIOR FILING DATE: 1998-11-17  
; PRIOR APPLICATION NUMBER: 60/108801  
; PRIOR FILING DATE: 1998-11-17  
; PRIOR APPLICATION NUMBER: 60/108849  
; PRIOR FILING DATE: 1998-11-18  
; PRIOR APPLICATION NUMBER: 60/112422  
; PRIOR FILING DATE: 1998-12-15  
; PRIOR APPLICATION NUMBER: 60/113296  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 60/113605  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113621  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/115558  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/115565  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/115733  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/119549  
; PRIOR FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 60/123618  
; PRIOR FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: 60/125259  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 60/125775  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: 60/126773  
; PRIOR FILING DATE: 1999-03-29  
; PRIOR APPLICATION NUMBER: 60/127887  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 60/130232  
; PRIOR FILING DATE: 1999-04-21  
; PRIOR APPLICATION NUMBER: 60/131022  
; PRIOR FILING DATE: 1999-04-26  
; PRIOR APPLICATION NUMBER: 60/131270  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/131291  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/131445  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: 60/134287  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/140650  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/140723  
; PRIOR FILING DATE: 1999-06-22  
;



## RESULTS AND DISCUSSION

;  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100848  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/100919  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/101477  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101738  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101741  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101786  
; PRIOR FILING DATE: 1998-09-25  
; PRIOR APPLICATION NUMBER: 60/101916  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101922  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/106178  
; PRIOR FILING DATE: 1998-10-28  
; PRIOR APPLICATION NUMBER: 60/106248  
; PRIOR FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: 60/106464  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 60/106905  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: 60/108787  
; PRIOR FILING DATE: 1998-11-17  
; PRIOR APPLICATION NUMBER: 60/108801  
; PRIOR FILING DATE: 1998-11-17  
; PRIOR APPLICATION NUMBER: 60/108849  
; PRIOR FILING DATE: 1998-11-18  
; PRIOR APPLICATION NUMBER: 60/112422  
; PRIOR FILING DATE: 1998-12-15  
; PRIOR APPLICATION NUMBER: 60/113296  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 60/113605  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113621  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/115558  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/115565  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/115733  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/119549  
; PRIOR FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 60/123618  
; PRIOR FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: 60/125259  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 60/125775  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: 60/126773  
; PRIOR FILING DATE: 1999-03-29  
; PRIOR APPLICATION NUMBER: 60/127887  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 60/130232  
; PRIOR FILING DATE: 1999-04-21  
; PRIOR APPLICATION NUMBER: 60/131022  
; PRIOR FILING DATE: 1999-04-26  
; PRIOR APPLICATION NUMBER: 60/131270  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/131291  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/131445  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: 60/134287  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/140650  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/140723  
; PRIOR FILING DATE: 1999-06-22

;  
; PRIOR APPLICATION NUMBER: 60/141037  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 60/144758  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/145698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: 60/146222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: 60/146963  
; PRIOR FILING DATE: 1999-08-03  
; PRIOR APPLICATION NUMBER: 60/149320  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/149638  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/151733  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/164418  
; PRIOR FILING DATE: 1999-11-09  
; PRIOR APPLICATION NUMBER: 60/166361  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR APPLICATION NUMBER: 60/169445  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 60/169495  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 60/169835

Alignment Scores:

Pred. No.:	3.87e-126	Length:	787
Score:	1524.50	Matches:	310
Percent Similarity:	58.89%	Conservative:	134
Best Local Similarity:	41.11%	Mismatches:	259
Query Match:	31.79%	Indels:	51
DB:	14	Gaps:	18

US-10-054-683-18 (1-2640) x US-10-230-163-90 (1-787)

QY	72	ATGTGGCTTGTGTTCTGTCTCAGCGGGCTCGGGGGCTGGGATGGAC---	AGTAATTTT	128
Db	4	LeuTrpLeu-----LeuLeuAlaGlyLeuCyGlyLeuAlaSerArgProGlyPhe	21	
QY	129	GATAGTTTACCTGTGCAAAATTACAGTTCCGGAGAAATACGGTCAATAATAAAGGAAGGA	188	
Db	22	GlnAsnSerLeuLeuGlnIleValIleProGluLysIleGlnThrAsnThrAsnAspSer	41	
QY	189	ATTGAA-----TCGCAGGCATCCTACAAAATTGTAATTCAAGGGAACCATATACT	239	
Db	42	SerGluIleGluTyrrGluGlnIleSerTyrrIleIleProIleAspGluLysLeuTyThr	61	
QY	240	GTGAATTTAATGCAAAAAAATTTTATCCCATATAATTTTAGAGTTTACAGTTTAGTGGC	299	
Db	62	ValHisLeuLysGlnArgTyrrPheLeuAlaAspAsnPheMetIleTyrrLeuTyrrAsn---	80	
QY	300	ACAGGAATTATGAAACCCTTGACCAAGATTTTCAGAAATTTCTGCCACTACCAAGGTAT	359	
Db	81	GlnGlySerMetAsnThrTyrrSerSerAspIleGlnThrGlnCysTyrrTyrrGlnGlyAsn	100	
QY	360	ATTGAAGGTTATCCAAAATCTGTGGTGATGGTAGCACATGTACTGTGACTCAGGGCGTA	419	
Db	101	IleGluGlyTyrrProaspSerMetValThrLeuSerThrCysSerGlyLeuArgGlyIle	120	
QY	420	CTACAGTTTGAAATCTTAGTTATGAATAGAACCCCTCGAGTCTTCAGTTGGCTTTGAA	479	
Db	121	LeuGlnPheGluAsnValSerTyrrGlyIleGluProLeuGluSerAlaValGluPheGln	140	
QY	480	CATGTAATTTACCAAGTAAACATAGAAAGCAGATGTTTCCTTATATATAGAGAGAT	539	
Db	141	HisValLeuTyrrLysLeuLysAsnGluAspAsnAspIleAlaIlePheIleAspArgSer	160	
QY	540	ATTGAATCAAGA-----GATCTGCTCTCTTAAATTTACAAAGCCAGACAGCCACAGCA	590	
Db	161	LeuLysGluGlnProMetaspAspAsnIlePheIleSerGluLysSerGluProAlaVal	180	
QY	591	GAT-----TTTCAAAGTATATAGAAATGTCATGTTATAGTTGAAAAACAATTGTATAAT	644	

Db 181 ProAspLeuPheProLeuTyrLeuGluMetHisIleValValAspLysThrLeuTyrAsp 200  
QY 645 CATATGGGCTGTGATACAACTGTTGCTCGCTCAAAAAGTTTCCAGTTCATTGGATTGACG 704  
Db 201 TyrTrpGlySerAspSerMetIleValThrAsnLysValIleGluIleValGlyLeuAla 220  
QY 705 AATGCTATTGTTGTTTCAATTAATTAACAATTATCTGCTCTCATTTGGAGCTTTGGATA 764  
Db 221 AsnSerMetPheThrGlnPheLysValThrIleValLeuSerSerLeuGluLeuTyrSer 240  
QY 765 GATGAAATAAATAATGCAACCACTGGAGAACTAATAGATTATTAACACATTTTAAAGA 824  
Db 241 AspGluAsnLysIleSerThrValGlyGluAlaAspGluLeuLeuGlnLysPheLeuGlu 260  
QY 825 TGGAAACATCTTATCTGTTTACGCTCATCATGTCGATTTTACTTGTGTACAGA 884  
Db 261 TrpLysGlnSerTyrLeuAsnLeuArgProHisAspIleAlaTyrLeuLeuIleTyrMet 280  
QY 885 GAAAGTCAAAATATGTTGGTGCACACCTTTCAAGGGAAGATGTGTATGCAAACTATGCA 944  
Db 281 AspTyrProArgTyrLeuGlyAlaValPheProGlyThrMetCysIleThrArgTyrSer 300  
QY 945 GGAGGTGTTGTCACCCAGCAACCAATAAGTCTGGAATCACTTGCAGTTATTTAGCT 1004  
Db 301 AlaGlyValAlaLeuTyrProLysGluIleThrLeuGluAlaPheAlaValIleValThr 320  
QY 1005 CAATTATTGAGCTTAGTATGGGATCACTTATGATGACATTAAACAATCCAGTGTCA 1064  
Db 321 GlnMetLeuAlaLeuSerLeuGlyIleSerTyrAspAspProLysLysCysGlnCysSer 340  
QY 1065 GGAGCTGCTGCTATTAATCAAGAACCAATTCATTTCAGTGTGTGAAGATCTTTAGT 1124  
Db 341 GluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValLysThrPheSer 360  
QY 1125 AACTGAGCTTCGAAGACTTGTGCACATTTATTTTCAAGCAGAAAGTCCAGTGTCTTAC 1184  
Db 361 SerCysSerLeuArgSerPheGlnAsnPheIleSerAsnValGlyValLysCysLeuGln 380  
QY 1185 AATCAGCTCGCTAGATCCTTTTCAACAGCAAGCAGTGTGTGAATGCAAGCTG 1244  
Db 381 AsnLysProGlnMetGlnLys---LysSerProLysProValCysGlyAsnGlyArgLeu 399  
QY 1245 GAAGCAGGAGGAGTGTGATGTGGGACTGAACAGGATTTGCTTATTTGGAGAACA 1304  
Db 400 GluGlyAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer 417  
QY 1305 TGCTGTGATATGGCATGTAGATTTAAAGCCGGTTCAACTGTGTGTGAAGCAGTGC 1364  
Db 418 CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyrLysGlyLeuCys 437  
QY 1365 TGGAAAACTGCTATTATTATCAAAAAGAAAGATGTGTAGGCT---TCCTTTGAAGAA 1421  
Db 438 CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu 457  
QY 1422 TGCAGCTCCCTGAATATTGCAATGGATTCATCTGCATCATGCCCCAGAAACCACTATGT 1481  
Db 458 CysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyProAspIleThrLeu 477  
QY 1482 CAGACTGGGCATCGTGTGAGTCAATCAATGGATCTGTATAGATGAGTGTGTATGAGT 1541  
Db 478 IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGlyAspCysHisAsp 497  
QY 1542 GGGATAAACAATGTCACACACATTTGGCAAGAAAGTAGATTTGGCCCTTCAGATGT 1601  
Db 498 LeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAlaProPheAlaCys 517  
QY 1602 TATTCTCACCTTAATCAAGACTGATGATCTCGAAACTGTGTGATAAGT---GATTCA 1658  
Db 518 TyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn 537  
QY 1659 GGATACACAGTGTGAAGCTGACAATCTGCAGTGGCGAAATTAATATGTAATATGTA 1718

Db 538 LysTyrValPheCysGlyTyrArgAsnLeuIleCysGlyArgLeuValCysThrTyrPro 557  
QY 1719 GGTAAATTTTTTATTACAAATTTCAAGAGCCACTATTATTATGCAACATAAGTGGACAT 1778  
Db 558 ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrAlaPheValArgAspSer 577  
QY 1779 CTCTGATCTTCTGCTGGAATTTGCGAGTGCATCATGACAGCCCAAGAGATGTGGATAAAA 1838  
Db 578 ValCysIleThrValAspTyrLysLeuProArgThrValProAspProLeuAlaValLys 597  
QY 1839 GATGGAACCTTCTGTTGTTCAATAAGTTTGCAGGAATCAAAAGATGTGTGAGTTCTTCA 1898  
Db 598 AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg 617  
QY 1899 TACTTG-----GGTTATGATTGTTACTACTGACAAATGCAATGATAGAGTGTATGC 1949  
Db 618 IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys 636  
QY 1950 AATAACAAAGACCTGCTCACTAGTGTCTTCAATTTTCTCCTCCAGATTGCTCAGTTCAA 2009  
Db 637 AspSerArgAsnLysCysHisCysSerProGlyTyrLysProAsnCysGlnIleArg 656  
QY 2010 TCAGATCTATGGCTGTGGAGTATT-----GACAGTGGCAATTTTCCACCT 2057  
Db 657 Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle----- 671  
QY 2058 GTAGCTATACAGCCAGACTCCTCTGAAAGCGCTACATTGAGAACATTTTACCATTCCAAA 2117  
Db 672 -----MetGluArgAlaSerGlyLysThr 679  
QY 2118 CCAATGAGATGG-----CCATTTTCTTATTTCATTCCTTTCTTTATTATTTCGTGTA 2171  
Db 680 GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal----- 696  
QY 2172 CTGATTGCTATAATGTGTGAAGTTAATTTCCAAAGGAAAAAATGGAGAACTAGGACTAT 2231  
Db 697 ---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu--- 714  
QY 2232 TCAAGCGATGAGCAACTGAAAGTGAAGTGAACCTAAAGGG 2273  
Db 715 -----GluGluPheProSerSerGluSerLysSerGluGly 726

RESULT 9  
US-10-230-338-90  
; Sequence 90, Application US/10230338  
; Publication No. US20030044934A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Gerritsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P35301C92  
; CURRENT APPLICATION NUMBER: US/10/230,338  
; CURRENT FILING DATE: 2002-08-28  
; PRIOR APPLICATION NUMBER: 10/119,480  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062287  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063549  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31

Db	221	AsnSerMetPheThrGlnPheIysValThrIleValLeuSerSerLeuGluLeuTrpSer	240
Qy	765	GATGAAATAAATTCGAACCACTGGAGAGAGCTAAATCAGTATTATACACACATTTTTTAAAGA	824
Db	241	AspGluAsnIysIleSerThrValGlyGluAlaAspGluLeuGlnIysPheLeuGlu	260
Qy	825	TGGAACAATCTTATCTTGTTTACGTCTCATGTATGTGGCATTTTTTACTGTTTACAGA	884
Db	261	TrpIysGlnSerTyrLeuAsnLeuAlaGProHieAspIleAlaTyrLeuLeuIleTyrMet	280
Qy	885	GAANAAGTCAAAATTATGTTGGTGTCAACCTTTCAGGGGAAGATGTGTGATGCAAACTATGCA	944
Db	281	AspTyrProArgTyrLeuGlyAlaValPheProGlyThrMetCysIleThrArgTyrSer	300
Qy	945	GGAGGTGTTCTTCTGACCCCGACCAACATAAAGTCTGGAATCACTTGCAGTATTTTAGCT	1004
Db	301	AlaGlyValAlaLeuTyrProIysGluIleThrLeuGluAlaPheAlaValIleValThr	320
Qy	1005	CAATTATTAGCCTTAGTATGGGATCACATTATGATGACATTAAACAAATGCCAGTGTCTCA	1064
Db	321	GlnMetLeuAlaLeuSerLeuGlyIleSerTyrAspAspProIysIysCysGlnCysSer	340
Qy	1065	GGAGCTCTGCATTATGATCCAGAGCAATTCATTTCAGTGGTGTGAGATCTTTAGT	1124
Db	341	GluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValIysThrPheSer	360
Qy	1125	AACTCGAGCTTCGAAGACTTTGCACATTTTATTTTCAAAGCAGAGAAGTCTCTTCAC	1184
Db	361	SerCysSerLeuArgSerPheGlnAsnPheIleSerAsnValGlyValIysCysLeuGln	380
Qy	1185	AATCAGCCTCGCTTAGATCCCTTTTTTCAAACAGCAGCAGTGTGTGGTAAATGCAAGCTG	1244
Db	381	AsnLysProGlnMetGlnIys---LysSerProLysProValCysGlyAsnGlyArgLeu	399
Qy	1245	GAAGCAGGAGAGAGTGTGACTGTGGACTGAACAGGATTTGTGCCCTTATTGGAGAAACA	1304
Db	400	GluGlyAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer	417
Qy	1305	TGCTGTGATATTGCCATCTAGATTAAAGCGGTTTCAAACCTGTCTGAAGACCATGC	1364
Db	418	CysCysAspPheArgThrCysValLeuIysAspGlyAlaIysCysTyrIysGlyLeuCys	437
Qy	1365	TGGCAAACTGTCTATTATGTCAAAAGAAAGAAATGTGTAGGCCT---TCCTTTGAAGAA	1421
Db	438	CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProIysAlaHieProGlu	457
Qy	1422	TGGACCTCCTCGAATATTGCAATGATCATCTGCATCATGCCAGAAACCATCATGTT	1481
Db	458	CysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyProAspIleThrLeu	477
Qy	1482	CAGACTGGGCATCCGTGTGGACTGAATCAATGATCTGTATAGATGGAGTTTGTATGAGT	1541
Db	478	IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGlyAspCysHieAsp	497
Qy	1542	GGGGATAAACAATGTACAGACACATTTGGCAAGAAGATAGAGTTTGGCCCTTCAGAAATGT	1601
Db	498	LeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAlaProPheAlaCys	517
Qy	1602	TATTCTCACCTTAATTCAAAGACTGATGATCTGGAAACCTGTGGTATAAGT---GATTCA	1658
Db	518	TyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn	537
Qy	1659	GGATACACACAGTGTGAGCTGACATCTGCTGCGGAAATTAATATATCTAAATATGTA	1718
Db	538	LysTyrValPheCysGlyTrpArgAsnLeuIleCysGlyArgLeuValCysThrTyrPro	557
Qy	1719	GGTAAATTTTATTACAAATTCGAAGAGCCACTTATTATTATGCCAACATAAGTGGACAT	1778
Db	558	ThrArgLysProPheHieGlnGluAsnGlyAspValIleTyrAlaPheValArgAspSer	577
Qy	1779	CTTCGATTTGCTGTGGAATTTGCCAGTGATCATGTCAGACAGCCAAAAGATGTGGATAAA	1838





Db 261 TrpLysGlnSerTyrLeuAsnLeuArgProHisAspIleAlaTyrLeuLeuIleTyrMet 280  
Qy 885 GAAAGTCAAAATTATGTTGGTCAACCTTTTCAAGGGAAGATGTGTGATGCAAACTATGCA 944  
Db 281 AspTyrProArgTyrLeuGlyAlaValPheProGlyThrMetCysIleThrArgTyrSer 300  
Qy 945 GGAGGTGTTGTCGACCCAGAACCAATAGTCTGGAATCAGTCTGGAGTATTTTAGCT 1004  
Db 301 AlaGlyValAlaLeuTyrProLysGluIleThrLeuGluAlaPheAlaValIleValThr 320  
Qy 1005 CAATTATTGAGCTTAGTATGGGATCATTATGATGACATTAACAATCCAGTGTCTCA 1064  
Db 321 GlnMetLeuAlaLeuSerLeuGlyIleSerTyrAspAspProLysLysCysGlnCysSer 340  
Qy 1065 GGAGCTCTCTGCAATTATGAATCCAGAAGCAATTCATTTCAGTGTGTGAAGATCTTTAGT 1124  
Db 341 GluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValLysThrPheSer 360  
Qy 1125 AACTGCACTTCGAGACTTTCACATTTTCAAGCAGAGAGTCCCGAGTGTCTTAC 1184  
Db 361 SerCysSerLeuArgSerPheGlnAsnPheIleSerAsnValGlyValLysCysLeuGln 380  
Qy 1185 AATCAGCTCGCTTAGATCCTTTTTCAAACAGCAGCAGTGTGTGTAATGCAAACTG 1244  
Db 381 AsnLysProGlnMetGlnLys---LysSerProLysProValCysGlyAsnGlyArgLeu 399  
Qy 1245 GAAGCAGGAGGAGTGTGAGTGTGGGACTGAACAGGATGTGCCCTTATTGGAGAACA 1304  
Db 400 GluGlyAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer 417  
Qy 1305 TGCTGTGATATGTCACATGTAGATTTAAAGCCGGTTCAAACTGTGTGTGAAGCACCATGC 1364  
Db 418 CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyrLysGlyLeuCys 437  
Qy 1365 TGGAAAACCTGCTATTATTATGTCAAAGAAGATGTGTAGGCT---TCTTTGAAGAA 1421  
Db 438 CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu 457  
Qy 1422 TGGACCTCCCTGAATATTGCAATGGATCATCTGCATCATGCCCAAGAACCACTATGTT 1481  
Db 458 CysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyProAspIleThrLeu 477  
Qy 1482 CAGACTGGGCATCGTGTGAGTGAATCAATGATCTGTATAGATGAGTGTGTATGAGT 1541  
Db 478 IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGlyAspCysHisAsp 497  
Qy 1542 GGGATAAACAATCTACAGACACATTTGGCAAGAGTGTGAGTGTGCCCTTCAGATGT 1601  
Db 498 LeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAlaProPheAlaCys 517  
Qy 1602 TATTCTCACCTTAATTCAAAGACTGTATCTGGAACCTGTGTATAGT---GATTCA 1658  
Db 518 TyrGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn 537  
Qy 1659 GGATACACAGTGTGAAGCTGCAATCTCAGTGGCGAAATTAATATGATTAATATGTA 1718  
Db 538 LysTyrValPheCysGlyTyrArgAsnLeuIleCysGlyArgLeuValCysThrTyrPro 557  
Qy 1719 GGTAAATTTTATACAAATTCAGAGCCACTATTATTTATGCAACATAAGTGGACAT 1778  
Db 558 ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrAlaPheValArgAspSer 577  
Qy 1779 CTCTGCATTCGTGGAAATTTGCCAGTGATCATGCAGACAGCCAAAGAGTGTGATAAAA 1838  
Db 578 ValCysIleThrValAspTyrLysLeuProArgThrValProAspProLeuAlaValLys 597  
Qy 1839 GATGGAATCTTCTGTGTTCAAAATAGGTTTGCAGGAATCAAGATGTGTGAGTCTTCA 1898  
Db 598 AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg 617  
Qy 1899 TACTTG-----GGTTATGATTGTACTACTGACAAATGCAATGATGATGAGTGTATGC 1949  
Db 1949

Db 618 IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys 636  
Qy 1950 AATAACAAAAGCACTGTACTGTAGTGTCTCATATTTACTCCAGATTTGCTCAGTTCAA 2009  
Db 637 AspSerArgAsnLysCysHisCysSerProGlyTyrLysProAsnCysGlnIleArg 656  
Qy 2010 TCAGATCTATGGCTGTGGTGGAGTATT-----GACAGTGGCAATTTTCCACCT 2057  
Db 657 Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle----- 671  
Qy 2058 GTAGTATACAGCCAGACCTCCCTGAAAGCGCTACATTCAGAACATTTACCATTCCAAA 2117  
Db 672 -----MetGluArgAlaSerGlyLysThr 679  
Qy 2118 CCAATGAGATGG-----CCATTTTCTTATTCATTCCTTCTTTATTTATTTCTGTGTA 2171  
Db 680 GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal----- 696  
Qy 2172 CTGATTCATATGTTGAAGTTAATTTCCAAAGGAAAAAATGGAGAACTGAGGACTAT 2231  
Db 697 ---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu--- 714  
Qy 2232 TCAAGCGATGACCAACTGAAAGTGAGAGTGAACCTAAAGCG 2273  
Db 715 -----GluIlePheProSerSerGluSerLysSerGluGly 726  
RESULT 11  
US-10-230-414-90  
; Sequence 90. Application US/10230414  
; Publication No. US20030050448A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Deenoyers, Luc  
; APPLICANT: Gerritsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3530P1C98  
; CURRENT APPLICATION NUMBER: US/10/230,414  
; CURRENT FILING DATE: 2002-08-28  
; PRIOR APPLICATION NUMBER: 10/119,480  
; PRIOR FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062287  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063549  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/069873  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 246  
; SEQ ID NO 90  
; LENGTH: 787  
; TYPE: PRT  
; ORGANISM: Homo Sapien





Db 341 GluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValLysThrPheSer 360  
QY 1125 AACTGCAGCTCGAAGACTTTTGCACATTTTATTTTCAAGACAGAGAGTCCAGTGTCTTAC 1184  
Db 361 SerCysSerLeuArgSerPheGlnAsnPhelSerAsnValGlyValLysCysLeuGln 380  
QY 1185 AATCAGCTCGCTTAGATCTCTTTTCAAAACAGCAGCAGTGTGTGTATGCAAGCTG 1244  
Db 381 AsnLysProGlnMetGlnLys---LysSerProLysProValCysGlyAsnGlyArgLeu 399  
QY 1245 GAGCAGAGAGAGTGTGACTGTGGAGCTGAACAGGATGTGCTTATTTGGAGAAACA 1304  
Db 400 GluGlyAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer 417  
QY 1305 TGCTGTGATATGTCACATAGATTTTAAAGCCGGTTCAAACTGTGCTGAGGACCATGC 1364  
Db 418 CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysThrLysGlyLeuCys 437  
QY 1365 TGGCAAAACTGTCTATTATTATGTCAAAAGAAAGATGTGTAGGCCT---TCTTTGAAAGAA 1421  
Db 438 CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu 457  
QY 1422 TGGACCTCCTGATATTGCAATGATGATCTGCATCATGCCAGAAACCATGATGTT 1481  
Db 458 CysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyProAspIleThrLeu 477  
QY 1482 CAGACTGGCATCGTGGACTCAATCAATGATGATGATGATGATGATGATGATGATGAT 1541  
Db 478 IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGlyAspCysHisAsp 497  
QY 1542 GGGGATAACAATGTACAGACACATTTGGCAAGAAAGATAGATTTTGGCCCTTCAGATGT 1601  
Db 498 LeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAlaProPheAlaCys 517  
QY 1602 TATTCTCACCTTAATTCAGACAGTGTATCTCGAACTGTGTGTATAAGT---GATTC 1658  
Db 518 TyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn 537  
QY 1659 GGATACACACAGTGTGAAGCTGACAATCTGAGCGGAAATTAATATGTAATATGTA 1718  
Db 538 LysTyrValPheCysGlyTyrArgAsnLeuIleCysGlyArgLeuValCysThrTyrPro 557  
QY 1719 GGTAAATTTTATTACAAATTCAGAGCCACTATTATTATGCAACATAAGTGGACAT 1778  
Db 558 ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrAlaPheValArgAspSer 577  
QY 1779 CTCTGCATTCGTGGAAATTTGCCAGTGCATGCACAGCCCAAGAGATGTGATGATAAA 1838  
Db 578 ValCysIleThrValAspTyrLysLeuProArgThrValProAspProLeuAlaValLys 597  
QY 1839 GATGGAATCTTGTGTTTCAATAAGTTTGCAGGAATCAAGAGATGTGTGAGTCTTCA 1898  
Db 598 AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg 617  
QY 1899 TACTTG-----GTTATGATTTGACTACAAATGCAATGCAATGATAGAGTGTATGC 1949  
Db 618 IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys 636  
QY 1950 AATAACAAAGACTGTCTACTGTAGTGTCTTATTTTACCTCAGATTCCTCAGTTCAA 2009  
Db 637 AspSerArgAsnLysCysHisCysSerProGlyTyrLysProAsnCysGlnIleArg 656  
QY 2010 TCAGATCTATGGCTGTGGAGTATT-----GACAGTGGCAATTTTCCACCT 2057  
Db 657 Ser-----LysGlyPheSerIlePheProGluAspMetGlySerIle----- 671  
QY 2058 GTAGCTATACAGCCAGACTCCCTGAAAGCGCTACATTCGAGAACATTTACCATTCAAA 2117  
Db 672 -----MetGluArgAlaSerGlyLysThr 679  
QY 2118 CCAATGATGAGG-----CCATTTTCTTATTCATTCCTTCTTATTTATTTCTGTGTA 2171

Db 680 GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal----- 696  
QY 2172 CTGATTGCTATAATGGTGAAGTAAATTTTCCAAAGGAAATAATGGAGAACTGAGGACTAT 2231  
Db 697 ---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu--- 714  
QY 2232 TCAAGCGATGAGCAACTGAACTGAGAGTGAGAACTTAAGGG 2273  
Db 715 -----GluGluPheProSerSerGluSerLysSerGluGly 726

RESULT 13  
US-10-216-159A-90  
; Sequence 90, Application US/10216159A  
; Publication No. US20030069397A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Gerritsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCES: P3530PIC6  
; CURRENT APPLICATION NUMBER: US/10/216,159A  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 10/119,480  
; PRIOR FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062287  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063549  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/069873  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 246  
; SEQ ID NO 90  
; LENGTH: 787  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-216-159A-90

Alignment Scores:  
Pred. No.: 3,87e-126 Length: 787  
Score: 1524.50 Matches: 310  
Percent Similarity: 58.89% Conservative: 134  
Best Local Similarity: 41.11% Mismatches: 259  
Query Match: 31.79% Indels: 51  
DB: 14 Gaps: 18

US-10-054-683-18 (1-2640) x US-10-216-159A-90 (1-787)

QY 72 ATGTGGTCTTGTCTTCTGCTCAGCGGCTCGCGGGCTCGGATGAC---AGTAATTTT 128  
Db 4 LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuAlaSerArgProGlyPhe 21

Qy	129	GATAGTTTACTGTGCGAAAATTACAGTTCCGCGAGAAAATACGGTCAATAATAAAGGAAGGA	188
Db	22	GlnAsnSerLeuLeuGlnIleValIleProGluLysIleGlnThrAnthrAsnAspSer	41
Qy	189	ATTGAA-----TCGCAAGGCATCCTCAAAATTGTAATTCGAAGGGAACCATATACT	239
Db	42	SerGluLeuGluTyrrGluGlnIleSerTyrrIleProIleAspGluLysLeuTyrrThr	61
Qy	240	GTGAATTTAATGCAAAAAAATCTTTTACCCTCAATAATTTAGAGTTTACAGTTATAGTGGC	299
Db	62	ValHisLeuLysGlnA-GTyrrPheLeuAlaAspAsnPheMetIleTyrrLeuTyrrAsn--	80
Qy	300	ACAGGAATTATGAACACCACTTGACCAAGATTTTCAGAAATTTTCGCCACTACCAAGGGTAT	359
Db	81	GlnGlySerMetAsnThrTyrrSerSerAspIleGlnThrGlnCysTyrrTyrrGlnGlyAsn	100
Qy	360	ATTGAAGGTTATCCAAAATCTGTGGTCATGCTGTAGCACATGTACTGGACACTCAGGGCGTA	419
Db	101	IleGluGlyTyrrProAspSerMetValThrLeuSerThrCysSerGlyLeuA-GGlyLe	120
Qy	420	CTACAGTTTGAAAATGTTAGTTATGGAATAGAACCCCTGGAGTCTTCAGTTGGCTTTGAA	479
Db	121	LeuGlnPheGluAsnValSerTyrrGlyIleGluProLeuGluSerAlaValGluPheGln	140
Qy	480	CATGTAATTTACCAAGTAAACATAGAAGCAGATGTTCTCTTATATATATAGAAGGAT	539
Db	141	HisValLeuTyrrLysLeuLysAsnGluAspAsnAspIleAlaIlePheIleAspArgSer	160
Qy	540	ATTGAATCAAGA-----GATCTGCTCTTAAATTACAAAGCGCAGAGCCACAGCAA	590
Db	161	LeuLysGluGlnProMetAspAspAsnIlePheIleSerGluLysSerGluProAlaVal	180
Qy	591	GAT-----TTTGCAAAATATATAGAATGATGTTATAGTTGAAAAACAATTTGATAAT	644
Db	181	ProAspLeuPheProLeuTyrrLeuGluMetHisIleValValAspLysThrLeuTyrrAsp	200
Qy	645	CATATGGGTCGTATACAACATGTTGTGCTCAAAAGTTTTCCAGTTGATGGATTGACG	704
Db	201	TyrrTrpGlySerAspSerMetIleValThrAsnLysValIleGluIleValGlyLeuAla	220
Qy	705	AATGCTATTTTGTTCATTTAAATATACAAATATTCTGTCTTCATTTGAGCTTTGGATA	764
Db	221	AsnSerMetPheThrGlnPheLysValThrIleValLeuSerSerLeuGluLeuTyrrPser	240
Qy	765	GATCAAAATAAAAATTGCAACCACTGGAGAGCTAATGAGTTATTACACACATTTTAAGA	824
Db	241	AspGluAsnLysIleSerThrValGlyLysAlaAspGluLeuGlnLysPheLeuGlu	260
Qy	825	TGGAACAACATCTTATCTTTTACGTCTCATGATGGCATTTTACTGTGTTTACAGA	884
Db	261	TrpLysGlnSerTyrrLeuAsnLeuArgProHisAspIleAlaTyrrLeuLeuIleTyrrMet	280
Qy	885	GAAGAAGTCAAAATTATGTTGTGCAACCTTTCAAGGGAAGATGTGTGATGCMAACTATGCA	944
Db	281	AspTyrrProArgTyrrLeuGlyAlaValPheProGlyThrMetCysIleThrArgTyrrSer	300
Qy	945	GGAGTGTGTGTCGACCCAGAACCATAGTCTGGAATCAGTTGAGTTATTTTACGT	1004
Db	301	AlaGlyValAlaLeuLeuTyrrProLysGluIleThrLeuGluAlaPheAlaValIleValThr	320
Qy	1005	CAATTATTGACCTTAGTATGGGATCACTTATGATGACATTAACAAATGCCAGTGTCTCA	1064
Db	321	GlnMetLeuAlaLeuSerLeuGlyIleSerTyrrAspAspProLysLysCysGlnCysSer	340
Qy	1065	GGAGCTGTCTGCATTATGAATCCAGAGCAATTCATTTCAGTGTGTGGAAGATCTTTAGT	1124
Db	341	GluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValLysThrPheSer	360
Qy	1125	AACGTGACGCTTCGAAGACTTTGCACATTTTATTTTCAAGCAGAGAAGTCCCAGTGTCTTCAC	1184
Db	361	SerCysSerLeuArgSerPheGlnAsnPheIleSerAsnValGlyValLysCysLeuGln	380
Qy	1185	AATCAGCCTCGCTTAGATCTCTTTTTCAAAACAGCAAGCAGTGTGTGTGTAATGCAAGCTG	1244

Db	381	AsnLysProGlnMetGlnLys---LysSerProLysProValCysGlyAsnGlyArgLeu	399
Qy	1245	GAAGCAGGACGAGGAGTGTGACTGTGGAGCATGAACAGGATTGTGCGCTTATTGGAGAAACA	1304
Db	400	GluGlyAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer	417
Qy	1305	TGCTGTGATATTGCCCATGTAGATTAAAGCCGGTTCAAACTGTGCTGTGAAGACCACTGTC	1364
Db	418	CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyrLysGlyLeuCys	437
Qy	1365	TGGCAAACTGCTCTATTATGTCAAAGAAGAATGTGTAGGCT--TCCTTTGAAGAA	1421
Db	438	CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu	457
Qy	1422	TGGACCTCCCTCAATATTCCAATTGGATCATCTGTCATCATGCCCCAGAAAAACCATGTGT	1481
Db	458	CysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyProAspIleThrLeu	477
Qy	1482	CAGACTGGGCATCCGTGTGGACTGAATCAATGAGTCTGTATAGATGGAGCTTTGTATGAGT	1541
Db	478	IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGlyAspCysHisAsp	497
Qy	1542	GGGGATAAACAATGTACAGACACATTTGGCAAGAAGTAGAGTTTGGCCCTTCAGAATGT	1601
Db	498	LeuAspAlaArgCysGluSerValPheGlyLysSerArgAsnAlaProPheAlaCys	517
Qy	1602	TATTCTCACCTTAATCAAAGACTGATGTATCTGGAACCTGTGGTATAAGT---GATTCA	1658
Db	518	TyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn	537
Qy	1659	GGATACACACAGTGTGAAGCTGACAATCTCGAGTCGGGAAAAATTAATATGTAATATGTA	1718
Db	538	LysTyrValPheCysGlyTyrArgAsnLeuIleCysGlyArgLeuValCysThrTyrPro	557
Qy	1719	GGTAAATTTTATACAAATTCCAAGAGCCACTATTATTATTCGCCACATAAGTGGACAT	1778
Db	558	ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrAlaPheValArgAspSer	577
Qy	1779	CTCTGCATTGCTGTGGAAATTTGCCAGTGCATCATGCAGACAGCCAAAAAGATGTGGATAAA	1838
Db	578	ValCysIleThrValAspTyrLysLeuProArgThrValProAspProLeuAlaValLys	597
Qy	1839	GATGGAATCTTCTGTGGTTCAAAATAAGTTTGCAGGAATCAAGATGTGTGAGTCTTCTCA	1898
Db	598	AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg	617
Qy	1899	TACTTG-----GGTTATGATTGTACTACTGCACAAATGCAATGATAGAGGTGATGC	1949
Db	618	IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys	636
Qy	1950	AATAACAAAAAGCACTCTCACTGTAGTGCTTCATATTACTTCAGATTGCTCAGTTTCAA	2009
Db	637	AspSerArgAsnLysCysHisCysSerProGlyTyrLysProProAsnCysGlnIleArg	656
Qy	2010	TCAGATCTATGGCTGTGGGAGTATT-----GACAGTGGCAATTTTCCACCT	2057
Db	657	Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle-----	671
Qy	2058	GTAGCTATACAGCCGACCTCCCTGAAGGGCTACATTGAGAACATTACCATTCCAAA	2117
Db	672	-----MetGluArgAlaSerGlyLysThr	679
Qy	2118	CCAATGAGATGG-----CCATTTTCTTATTCATCTCTTCTTATTTATTTCTGTGTA	2171
Db	680	GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuVal-----	696
Qy	2172	CTGATTGCTATAATGTGGAAGTTTAATTTCCAAAGGAAAAAATGGAGAACTGAGGACTAT	2231
Db	697	---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu---	714
Qy	2232	TCAAGCGATGAGCAACTGTAAGATGAGGTGAACCTAAAGGG	2273

Db 715 -----GluGluPheProSerSerGluSerLysSerGluGly 726

## RESULT 14

US-10-218-849-90

; Sequence 90, Application US/10218849  
; Publication No. US20030073814A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Deanovers, Luc  
; APPLICANT: Gerritsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Guney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3530P1C11  
; CURRENT APPLICATION NUMBER: US/10/218,849  
; CURRENT FILING DATE: 2002-08-12  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 246  
; SEQ ID NO 90  
; LENGTH: 787  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-218-849-90

## Alignment Scores:

Pred. No.:	3,87e-126	Length:	787
Score:	1524.50	Matches:	310
Percent Similarity:	58.89%	Conservative:	134
Best Local Similarity:	41.11%	Mismatches:	259
Query Match:	31.79%	Indels:	51
DB:	14	Gaps:	18

US-10-054-683-18 (1-2640) x US-10-218-849-90 (1-787)

Qy	72	ATGTGGCTGTTTCTGCTCAGCGGCTCGCGGCTGGGATGGAC---AGTAATTTT	128
Db	4	LeuTriPhe-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe	21
Qy	129	GATAGTTTACCTGTCGCAATACAGTTCGCGAGAAATACGGTCAATAATAAGGAGGA	188
Db	22	GlnAsnSerLeuLeuGlnIleValIleProGluLysIleGlnThrAsnAspSer	41
Qy	189	ATTGAA-----TCGCGGCTCCTACAAATTTGTAATTTGAAGGAAACCATATACT	239
Db	42	SerGluIleGluTyrGluGlnIleSerTyrIleIleProIleAspGluLysLeuTyrThr	61
Qy	240	GTGAATTAATGCAAAATAATTTTACCCCATTAATTTAGAGTTTACAGTTATAGTGC	299
Db	62	ValHisLeuLysGlnArgTyrPheLeuAlaAspAsnPheMetIleTyrLeuTyrAsn---	80
Qy	300	ACAGGATTTAGAACCACTTGACCAAGATTTTCAGAAATTTCTGCCACTACCAAGGTAT	359
Db	81	GlnGlySerMetAsnThrTyrSerSerAspIleGlnThrGlnCysTyrTyrGlnGlyAsn	100
Qy	360	ATTGAAGTTTATCAAAATCTGTGGTATGTTAGCACATGTTACTGGACTCAGGGCGGTA	419
Db	101	IleGluGlyTyrProAspSerMetValThrLeuSerThrCysSerGlyLeuArgGlyIle	120
Qy	420	CTACAGTTTGAATGTTAGTATGGAATGAAACCCCTGAGATCTTCAGTTGGCTTTGAA	479
Db	121	LeuGlnPheGluAsnValSerTyrGlyIleGluProLeuGluSerAlaValGluPheGln	140
Qy	480	CATGTAATTTACCAAGTAAACATAAGAACGACAGATGTTTCCTTATATATAGNAGGAT	539
Db	141	HisValLeuTyrLysLeuLysAsnGluAspAsnAspIleAlaIlePheIleAspArgSer	160

Qy	540	ATTGAATCAAGA-----GATCTCTCCTTTAAATTAACAAGCGCAGACCCACGACAA	590
Db	161	LeuLysGluGlnProMetAspAsnIlePheIleSerGluLysSerGluProAlaVal	180
Qy	591	GAT-----TTTGCAAGTATATAGAAATCATGTTATAGTTGAAACAAATTTGTAAT	644
Db	181	ProAspLeuPheProLeuTyrLeuGluMetHisIleValValAspLysThrLeuTyrAsp	200
Qy	645	CATATGGGCTCTGATCAACTGTTGCGCTCAAAAGTGTTCACAGTTGATGGATGACG	704
Db	201	TyrTrpGlySerAspSerMetIleValThrAsnLysValIleGluIleValGlyLeuAla	220
Qy	705	AATGCTATTTTCTTTTCATTAATATTAACAATATTCTGCTCTTCATGGAGCTTTGATA	764
Db	221	AsnSerMetPheThrGlnPheLysValThrIleValLeuSerSerLeuGluLeuTrpSer	240
Qy	765	GATGAAAAATAAATTCGACCACTGGAGAGCTTAATAGATTATACACACATTTTAAAG	824
Db	241	AspGluAsnLysIleSerThrValGlyGluAlaAspGluLeuGlnLysPheLeuGlu	260
Qy	825	TGGAAACACATCTTCTGTTTACGTCCTCATGATGTCGCTTTTACTGTTTACGACA	884
Db	261	TrpLysGlnSerTyrLeuAsnLeuArgProHisAspIleAlaTyrLeuLeuIleTyrMet	280
Qy	885	GAAGAAGTCAAAATTTATGTTGGTCAACTTTTCAAGGGAAGATGCTGATGCAAACTATGCA	944
Db	281	AspTyrProArgTyrLeuGlyAlaValPheProGlyThrMetCysIleThrArgTyrSer	300
Qy	945	GGAGGTGTTGTTCTGCACCCGACCAATAGTCTGGAATCACTTGCGAGTTATTTAGCT	1004
Db	301	AlaGlyValAlaLeuTyrProLysGluIleThrLeuGluAlaPheAlaValIleValThr	320
Qy	1005	CAATTTATGAGCCTTAGTATGGGATCACTTATGATGACATTAAACAAATCCAGTGTCTCA	1064
Db	321	GlnMetLeuAlaLeuSerLeuGlyIleSerTyrAspAspProLysLysCysGlnCysSer	340
Qy	1065	GGAGCTGCTGTCATTATGAATCCAGAGCAATTCATTTCAAGTGTGTGAAGATCTTTAGT	1124
Db	341	GluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValLysThrPheSer	360
Qy	1125	AATGAGCTTTCGAAAGCTTTGACATTTATTTTAAAGCAGAGTCCCGAGTCTTTCAC	1184
Db	361	SerCysSerLeuArgSerPheGlnAsnPheIleSerAsnValGlyValLysCysLeuGln	380
Qy	1185	AATCAGCTCGCTTACATCCTTTTCAAAACACCAAGCAGTGTGTGTAATGCAAGCTG	1244
Db	381	AsnLysProGlnMetGlnLys---LysSerProLysProValCysGlyAsnGlyArgLeu	399
Qy	1245	GAAGCAGGAGAGGAGTGTGACTGTGGGACTGAACAGAGTTTGTGCCCTTATTGGAGAAACA	1304
Db	400	GluLysAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer	417
Qy	1305	TGCTGTGATATGCCCATGTAGATTTAAAGCCGGTTCAAACTGTGCTGAGAGCAACATGC	1364
Db	418	CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyrLysGlyLeuCys	437
Qy	1365	TGCGAAACCTCTCTTTATGTCAAAGAAGAATGTGTAGGCT---TCCTTTTGAAGA	1421
Db	438	CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu	457
Qy	1422	TGCGACTCTCCTGAATTTGCAATGATCATCTGCAATCATGCCCCGAAACACCATATGTT	1481
Db	458	CysAspIleAlaGluAsnCysAsnGlySerSerProGluCysGlyProAspIleThrLeu	477
Qy	1482	CAGACTGGGATCCGTGGTGGATCAATCAATGATCTGATATAGATGAGTTTGTATGAGT	1541
Db	478	IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrAspGlyAspCysHisAsp	497
Qy	1542	GGGATTAACAATGTCACACACATTTTGGCAAGAAGTAGAGTTTGGCCCTTCCAGATGT	1601
Db	498	LeuAspAlaArgCysGluSerValPheGlyLysGlySerArgAsnAlaProPheAlaCys	517
Qy	1602	TATTTCTCACCTTAATTTCAAGAGACTGATGATCTCGAAACTGTGGTATAGT---GATTCA	1658

Db 518 TyGluLulleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn 537  
Qy 1659 GGATACACACAGTGTGAAGCTGCAATCTCGAGTGGCGAAATTAATATATGTA 1718  
Db 538 LysTyValPheCysGlyTrpArgAsnLeuLeuCysGlyArgLeuValCysThrTyPro 557  
Qy 1719 GGTAAATTTTATACAAATTCGAAGCCACTATTATATGCAACATAAGTGGACAT 1778  
Db 558 ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyAlaPheValArgAspSer 577  
Qy 1779 CTCTGCAATCTGTGGAATTTGCCAGTGATCATCCAGACAGCCAAAGATGTGATATAA 1838  
Db 578 ValCysIleThrValAspTyTrpLysLeuProArgThrValProAspProLeuAlaValLys 597  
Qy 1839 GATGGAATCTTCTGTGTTCAATAAAGTTTGAGGAAATCAAGATGTGTGAGTTCTTCA 1898  
Db 598 AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg 617  
Qy 1899 TACTTG-----GGTTAGATTGTACTGACAAATGCAATGCAATGATAGAGGTGATGC 1949  
Db 618 IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys 636  
Qy 1950 AATAACAAAGACACTGCTACTGTAGTCTTCAATTTTACCTCCAGATTCCTCAGTTCAA 2009  
Db 637 AspSerArgAsnLysCysHisCysSerProGlyTyTrpLysProAsnCysGlnIleArg 656  
Qy 2010 TCAGATCTATGGCTGTGGGAGTATT-----GACAGTGGCAATTTTCCACCT 2057  
Db 657 Ser-----LysGlyPheSerIlePheProGluLysMetGlySerIle----- 671  
Qy 2058 GTAGTATACAGCCAGACTCCCTGAAAGCGGTACATTCGAGAACATTTACCATTCCAA 2117  
Db 672 -----MetGluArgAlaSerGlyLysThr 679  
Qy 2118 CCAATGAGATG-----CCATTTTCTTATTATTCCTTCTTTATTATTTCTGTGTA 2171  
Db 680 GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal----- 696  
Qy 2172 CTGATGCTATATGTTGGAAGTTAATTTCCAAAGGAAAAATGGAGAACTGAGGACTAT 2231  
Db 697 ---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu--- 714  
Qy 2232 TCAAGCGATGAGCAACTGAAAGTGAGTGAGTGAACCTTAAGG 2273  
Db 715 -----GluGluPheProSerSerGluSerLysSerGluGly 726  
RESULT 15  
US-10-227-873-90  
; Sequence 90, Application US/10227873  
; Publication No. US20030073816A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Deenoyers, Luc  
; APPLICANT: Gerritsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: F3530FIC72  
; CURRENT APPLICATION NUMBER: US/10/227,873  
; CURRENT FILING DATE: 2002-08-26  
; PRIOR FILING DATE: 10/119,480  
; PRIOR FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062287

; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063549  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/069873  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082804  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/084441  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/085323  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/086392  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089905  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090691  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/095302  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/095318  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/095916  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: 60/096146  
; PRIOR FILING DATE: 1998-08-11  
; PRIOR APPLICATION NUMBER: 60/096791  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 60/097986  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: 60/098544  
; PRIOR FILING DATE: 1998-08-31  
; PRIOR APPLICATION NUMBER: 60/099596  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099598  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099803  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099811  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099812  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099816  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/100038  
; PRIOR FILING DATE: 1998-09-11

Qy	128	ATGTGGGCTCTGTTCTGCTCAGCGGGCTCGCGCGGCTCGCGATGGAC---	AGTAATTTT
Db	128	LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe	21
Qy	168	GATAGTTTACCTGTGCAATACAGTTCGGGAGAAATACGGTCAATAATAAAGGAGGA	188
Db	168	LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe	21
Qy	239	ATTGAA-----TCGCAGCGCATCTCTACAAAATTGTAATTGAAGGAGAAACCATATACT	239
Db	239	LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe	21
Qy	299	GTCGAATTTAATCCAAAAAATCTTTTACCATTAATTTTAGAGTTTACAGTTTATAGTGGC	299
Db	299	LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe	21
Qy	359	ACAGGAATTTATCAAAACCACCTTGACCAAGATTTTCAGAAATTTCTCCACTACCAAGGGTAT	359
Db	359	LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe	21
Qy	419	ATTGAAGGTTATCCAAAATCTGTGTGTAGTTAGCATTGTACTCGACTCAGGGCGGTA	419
Db	419	LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe	21
Qy	479	CTACAGTTTGAAAATGTTTAGTTATGGAATACAAACCTCGAGTCTTCCAGTTGGCTTTGAA	479
Db	479	LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe	21
Qy	539	CATGTAATTTACCAAGTAAAAACATAAGAAAGCAGATGTTTCTCTTATATAATGAGAAGAT	539
Db	539	LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe	21
Qy	160	HisValLeuTyrlsLeuLysLeuValAsnGluAspAsnAspIleAlaIlePheIleAspArgSer	160
Db	160	LeuTrpLeu-----LeuLeuAlaGlyLeuCysGlyLeuLeuAlaSerArgProGlyPhe	21



```

QY 540 ATTGAATCAAGA-----GATCTGCTCCTTTAAATTACAAAGCGCAGAGCCACAGCAA 590
Db 161 LeuylsGluInProMetAspAspAsnIlePheIleSerGluLysSerGluProAlaVal 180
QY 591 GAT-----TTTCAAAAGTATATAGAAATCATGTTATATCTTGAATAAACAATTTGTAAT 644
Db 181 ProAspLeuPheProLeuTyrlsLeuGluMetHisIleValValAspLysThrLeuTyrlsAsp 200
QY 645 CATATGGGGTCTGATACAACTGTTGTGCTCAAAAAGTCTTTCAGTGTGATTGGATGACG 704
Db 201 TyrlsGlySerAspSerMetIleValThrAsnLysValIleGluIleValGlyLeuAla 220
QY 705 AATGCTATTTTGTTCATTTATATATACAAATTTCTGCTCTCATTTGGAGCTTTGGATA 764
Db 221 AsnSerMetPheThrGlnPheLysValThrIleValLeuSerSerLeuGluLeuTrpSer 240
QY 765 GATTGAAATAAATTCGAAACCACTGGAGAAAGCTAAATGAGTTATTACACACATTTTAAAGA 824
Db 241 AspGluAsnLysIleSerThrValGlyGluAlaAspGluLeuGlnLysPheLeuGlu 260
QY 825 TGGAAAACATCTTATCTTGTGTTTACGTCCTCATGATGTGCATTTTACTTGTTTACAGA 884
Db 261 TrpLysGlnSerTyrlsLeuAsnLeuArgProHisAspIleAlaTyrlsLeuLeuIleTyrlsMet 280
QY 885 GAAAGTCAAAATTTATGTTGTGTGCAACCTTTCAAGGGAGATGTGTGATGCAAACTATGCA 944
Db 281 AspTyrlsProArgTyrlsLeuGlyAlaValPheProGlyThrMetCysIleThrArgTyrlsSer 300
QY 945 GGAGGTGTTCTTCTGCAACCCAGAACCAATAGCTGGAATCACTTGCAGTTATTATAGCT 1004
Db 301 AlaGlyValAlaLeuTyrlsProLysGluIleThrLeuGluAlaPheAlaValIleValThr 320
QY 1005 CAATTTATTGAGCCTTAGTATGGGGATCACATTTATGATGACATTAACAAATCCCAAGTGTCA 1064
Db 321 GlnMetLeuAlaLeuSerLeuGlyIleSerTyrlsAspAspProLysLysCysGlnCysSer 340
QY 1065 GGAGCTGCTGCAATATGATATCAGAAAGCAATTCATTTAGTGTGTGAGATCTTTAGT 1124
Db 341 GluSerThrCysIleMetAsnProGluValValGlnSerAsnGlyValLysThrPheSer 360
QY 1125 AACTGAGACTTCGAAGACTTTGCACATTTTATTTCAAGCAGAGTCCCAAGTCTCTTCAC 1184
Db 361 SerCysSerLeuArgSerPheGlnAsnPheIleSerAsnValGlyValLysCysLeuGln 380
QY 1185 AATCAGCCTCGCTTAGATCCTTTTTTCAAAACAGCAAGCAGTGTGTGTAATGCAAAAGCTG 1244
Db 381 AsnLysProGlnMetGlnLys---LysSerProLysProValCysGlyAsnGlyArgLeu 399
QY 1245 GAAGCAGGAGGAGGTGATGCTGTGGACTGAACAGAGATTGTGCCCTTATTGGAGAAACA 1304
Db 400 GluGlyAsnGluIleCysAspCysGlyThrGluAlaGlnCysGly-----ProAlaSer 417
QY 1305 TGCTGTGATATTGCCATGATAGATTAAAGCCGTTCAAACTGTGCTGAAGCAACCATGC 1364
Db 418 CysCysAspPheArgThrCysValLeuLysAspGlyAlaLysCysTyrlsGlyLeuCys 437
QY 1365 TGGCAAACTGCTATTATTATGTCAAAAGAAAGAAATGTGTAGGCT---TCTTTTGAAGAA 1421
Db 438 CysLysAspCysGlnIleLeuGlnSerGlyValGluCysArgProLysAlaHisProGlu 457
QY 1422 TGGCACTCCTCGATATTCGAATTCGATCATCTGATCATGCCAGAAACCACTATGTT 1481
Db 458 CysAspIleAlaGluAsnCysAsnGlySerProGluCysGlyProAspIleThrLeu 477
QY 1482 CAGACTGGGCATCCGTGTGGACTGAAATCAATGATCTGTATAGATGGAGTTTGTATGAGT 1541
Db 478 IleAsnGlyLeuSerCysLysAsnAsnLysPheIleCysTyrlsAspGlyAspCysHisAsp 497
QY 1542 GGGGATAAACAATGTCACAGACACATTTGGCAAGAAAGTAGAGTTTGGCCCTTCAGAAATGT 1601
Db 498 LeuAspAlaArgCysGluSerValPheGlyIysGlySerArgAsnAlaProPheAlaCys 517
QY 1602 TATTCTCACCTTAATCAAGACTGATGTATCTCGAAACTGTGTATAGT---GATTCA 1658

```

```

Db 518 TyrGluGluIleGlnSerGlnSerAspArgPheGlyAsnCysGlyArgAspArgAsnAsn 537
QY 1659 GGATACACACAGTGTGAAGCTGACAATTCGACGTGGGAAAAATTAATATATTAATATGTA 1718
Db 538 LysTyrlsValPheCysGlyTrpArgAsnLeuIleCysGlyArgLeuValCysThrTyrlsPro 557
QY 1719 GGTAAATTTTATTACAAATTCCAAGAGCCACTATTATTATCCCAACATAAAGTGGACAT 1778
Db 558 ThrArgLysProPheHisGlnGluAsnGlyAspValIleTyrlsAlaPheValArgAspSer 577
QY 1779 CTCTGCATTCCTGTGGAATTTGGCAGTGATCATGCAGACAGCCAAAGATGTCGATGATAA 1838
Db 578 ValCysIleThrValAspTyrlsLysLeuProArgThrValProAspProLeuAlaValLys 597
QY 1839 GATGGAATCTTGTGTGTTCAATAAGTGTTCGAGGAATCAAAAGATGTGTGAGTTCTTCA 1898
Db 598 AsnGlySerGlnCysAspIleGlyArgValCysValAsnArgGluCysValGluSerArg 617
QY 1899 TACTTG-----GGTTATGATTGTACTACTGACAAATGCAATGCAATGATAGAGTGTATGC 1949
Db 618 IleIleLysAlaSerAlaHisValCys---SerGlnGlnCysSerGlyHisGlyValCys 636
QY 1950 AATAACAAAAGCACTGTCACTGTGCTTTCATATTACCTCCAGATTCCTCAGTTCAA 2009
Db 637 AspSerArgAsnLysCysHisCysSerProGlyTyrlsProAsnCysGlnIleArg 656
QY 2010 TCAGATCTATGGCTGTGTGGAGTATT-----GACAGTGGCAATTTTCCACCT 2057
Db 657 Ser-----LysGlyPheSerIlePheProGluGluAspMetGlySerIle----- 671
QY 2058 GTAGCTATACCAAGCCAGACTCCCTGAAAGCGCTACATTGAGAACATTTTACCATTCCAA 2117
Db 672 -----MetGluArgAlaSerGlyLysThr 679
QY 2118 CCAATGAGATGG-----CCATTTTCTTATTCATTCTTCTTCTTCTTCTTCTTCTGTA 2171
Db 680 GluAsnThrTrpLeuLeuGlyPheLeuIleAlaLeuProIleLeuIleVal----- 696
QY 2172 CTGATTGCTATATGTTGAAGTTAATTTCCAAAGGAAAAAATGGAGAACTGGAGGACTAT 2231
Db 697 ---ThrThrAlaIleValLeuAlaArgLysGlnLeuLysLysTrpPheAlaLysGlu--- 714
QY 2232 TCAAGCGATGAGCAACCTGAAAGTGAAGTGAAGTGAACCTAAAGGG 2273
Db 715 -----GluGluPheProSerSerGluSerLysSerGluGly 726

```

Search completed: January 10, 2005, 22:20:41  
Job time : 396.5 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 21:47:51 ; Search time 160 Seconds  
(without alignments)  
1645.671 Million cell updates/sec

Title: US-10-054-683-19

Perfect score: 3984

Sequence: 1 MWVFLSLGLGLRMSNFD.....WRTEYSSDEQPESESEPKG 734

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_23Sep04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3984	100.0	734	2 AAR87037	Aar87037 Human PH3
2	3984	100.0	734	6 ABJ19246	Abj19246 Human can
3	3984	100.0	734	7 ADB75186	Adb75186 Prostate
4	3969.5	99.6	735	7 ADB75184	Adb75184 Prostate
5	3898	97.8	753	7 ADB75182	Adb75182 Prostate
6	3551	89.1	651	2 AAR87034	Aar87034 Human PH3
7	2440	61.2	735	2 AAR87036	Aar87036 Mouse PH3
8	1524.5	38.3	787	5 AAU83636	Aau83636 Human PRO
9	1524.5	38.3	787	6 ABU80783	Abu80783 Human PRO
10	1524.5	38.3	787	6 ABO33749	Abo33749 Novel hum
11	1524.5	38.3	787	6 ABU82092	Abu82092 Novel hum
12	1524.5	38.3	787	6 ABJ72272	Abj72272 Human PRO
13	1524.5	38.3	787	6 ABJ72400	Abj72400 Human PRO
14	1524.5	38.3	787	6 ABO34295	Abo34295 Human sec
15	1524.5	38.3	787	7 ABJ72102	Abj72102 Human mem
16	1524.5	38.3	787	7 ADB83580	Adb83580 Novel hum
17	1524.5	38.3	787	7 ADB80686	Adb80686 Novel hum
18	1524.5	38.3	787	7 ADB73227	Adb73227 Novel hum
19	1524.5	38.3	787	7 ADB78309	Adb78309 Novel hum
20	1524.5	38.3	787	7 ADB84957	Adb84957 Human PRO
21	1524.5	38.3	787	7 ADB78063	Adb78063 Novel hum
22	1524.5	38.3	787	7 ADB87129	Adb87129 Human PRO
23	1524.5	38.3	787	7 ADB84711	Adb84711 Human PRO
24	1524.5	38.3	787	7 ADB83826	Adb83826 Novel hum
25	1524.5	38.3	787	7 ADB72981	Adb72981 Novel hum

Adc36819 Human PRO  
Adc21809 Human PRO  
Adc49840 Novel hum  
Adc49039 Novel hum  
Adc49556 Novel hum  
Adc47417 Novel hum  
Adc47162 Novel hum  
Adc78037 Novel hum  
Add06272 Novel hum  
Adc77791 Novel hum  
Add50754 Novel hum  
Add51000 Novel hum  
Add50481 Human PRO  
Add50235 Human PRO  
Add51246 Novel hum  
Adc48793 Novel hum  
Ade20964 Novel hum  
Ade05808 Human PRO  
Add75037 Human PRO  
Add75783 Novel hum

ALIGNMENTS

RESULT 1  
AAR87037  
ID AAR87037 standard; protein; 734 AA.  
XX AAR87037;  
DT 30-SEP-1996 (first entry)  
XX Human PH30 beta chain sperm protein.  
DE Human PH30 beta chain sperm protein.  
XX Human PH30 beta chain sperm protein; contraceptive;  
KW FEE integrin binding domain.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Region 84..734  
FT /note= "see AAR87035"  
XX  
PN WO9535118-A1.  
PD 28-DEC-1995.  
XX  
PF 06-JUN-1995; 95WO-US007295.  
XX  
PR 20-JUN-1994; 94US-00264101.  
XX  
PA (MERI ) MERCK & CO INC.  
XX  
PI Alves K, Gupta SK, Hollis GF;  
XX WPI; 1996-058212/06.  
XX N-PSDB; AAT07328.  
XX  
PT Human and mouse sperm protein PH30 beta chain and related DNA - useful in  
XX contraceptive vaccines.  
XX  
PS Example 2; Page 45-48; 85pp; English.  
XX  
CC Human PH30 beta chain sperm protein, having a FEE integrin binding  
CC domain, is 58.9% identical to mouse and 56.5% identical to guinea pig  
CC PH30 beta. The protein may be produced recombinantly and used in a  
CC contraceptive composition containing an effective adjuvant and an amount  
CC of sperm protein which is effective for the stimulation of antibodies  
CC which bind to sperm protein in vivo, thereby preventing or substantially  
XX reducing the rate of sperm-egg fusion  
XX Sequence 734 AA;

```
Query Match      100.0%; Score 3984; DB 2; Length 734;
Best Local Similarity 100.0%; Pred. No. 5.5e-292;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWVFLSLGGLRMDNSFSLPQVITVPEKIRIIKEGIESQASYKIVIEGKPYTVNLM 60
Db 1 MWVFLSLGGLRMDNSFSLPQVITVPEKIRIIKEGIESQASYKIVIEGKPYTVNLM 60
Qy 61 QKNFLPHNFRVYSYSGTGMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGLQFE 120
Db 61 QKNFLPHNFRVYSYSGTGMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGLQFE 120
Qy 121 NVSYGIEPLESSVGFHEVYQVHKKADVSLYNEKDIESRDLSFKLOSAPQDFAKYIE 180
Db 121 NVSYGIEPLESSVGFHEVYQVHKKADVSLYNEKDIESRDLSFKLOSAPQDFAKYIE 180
Qy 181 MHVIVEKQLYNHMGSDTTVAQVQLIGLTNAIFVSNITIIILSSLELWIDENKIATTG 240
Db 181 MHVIVEKQLYNHMGSDTTVAQVQLIGLTNAIFVSNITIIILSSLELWIDENKIATTG 240
Qy 241 EANELLHTFLRWKTSYLVLRPHDVAFLLVYREKSNYVGATFQGMKCDANVAGGVVLHPRT 300
Db 241 EANELLHTFLRWKTSYLVLRPHDVAFLLVYREKSNYVGATFQGMKCDANVAGGVVLHPRT 300
Qy 301 ISLESLAVILAQLLSLMSGITYDDINKCQCGSAGVCIWNPPEAIHFSGVKIFSNCSFEDFAH 360
Db 301 ISLESLAVILAQLLSLMSGITYDDINKCQCGSAGVCIWNPPEAIHFSGVKIFSNCSFEDFAH 360
Qy 361 FISKQKSOCLHNPRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420
Db 361 FISKQKSOCLHNPRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420
Qy 421 KAGSNCAEGPCCNCLFMSKERMCRPSFECDLPEYCNCGSSASCPENHYVOTGHPCGLNQ 480
Db 421 KAGSNCAEGPCCNCLFMSKERMCRPSFECDLPEYCNCGSSASCPENHYVOTGHPCGLNQ 480
Qy 481 WICIDGCMGSDKQCTDTFGKEVEFGPSECSYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540
Db 481 WICIDGCMGSDKQCTDTFGKEVEFGPSECSYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540
Qy 541 CGKLIKCYVGKFLQIPRATIIYANISGHLCIAVEFADSHADSKQMIKDGTSQSGNKVC 600
Db 541 CGKLIKCYVGKFLQIPRATIIYANISGHLCIAVEFADSHADSKQMIKDGTSQSGNKVC 600
Qy 601 RNQRCVSSSYLYGDTTDCNDRGVCNKKHCHCSASYLPDCSVQSDLPWGGSIDSGNF 660
Db 601 RNQRCVSSSYLYGDTTDCNDRGVCNKKHCHCSASYLPDCSVQSDLPWGGSIDSGNF 660
Qy 661 PPVAIPARLPERRYENIYHSKPMRWPFLLPFPFFIIFCVLIAIMVKVNFQKKWRTEDY 720
Db 661 PPVAIPARLPERRYENIYHSKPMRWPFLLPFPFFIIFCVLIAIMVKVNFQKKWRTEDY 720
Qy 721 SSDEQPESESEPKG 734
Db 721 SSDEQPESESEPKG 734

RESULT 2
ABU19246
ID ABU19246 standard; protein; 734 AA.
XX
AC ABU19246;
XX
DT 28-MAR-2003 (first entry)
XX
DE Human cancer/testis antigen - SEQ ID No 19.
XX
KW Human; gene therapy; vaccine; cancer; cancer/testis antigen; CT antigen.
XX
OS Homo sapiens.
XX
PN W0200278526-A2.
```

Db 481 WICIDGVCMSGDKQCTDTTFCGEVFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540  
Qy 541 CGKLICKYVGKFLLOIPRATIIYANISGHLCIAVEFASDHADSQKMWIKDGTSCGSKNVC 600  
Db 541 CGKLICKYVGKFLLOIPRATIIYANISGHLCIAVEFASDHADSQKMWIKDGTSCGSKNVC 600  
Qy 601 RNORCVSSSYLGYDCTTDCNDRGVNCKKHCHCSASYLPPDCSVQSDLPWGSIDSGNF 660  
Db 601 RNORCVSSSYLGYDCTTDCNDRGVNCKKHCHCSASYLPPDCSVQSDLPWGSIDSGNF 660  
Qy 661 PPVAIPARLPERRYENIYHSPKMRWPFLLFIPFFIFCVLIAIMVKVNFQKKWRTEDY 720  
Db 661 PPVAIPARLPERRYENIYHSPKMRWPFLLFIPFFIFCVLIAIMVKVNFQKKWRTEDY 720  
Qy 721 SSDEQPESESEPKG 734  
Db 721 SSDEQPESESEPKG 734

RESULT 3  
ID ADB75186 standard; protein; 734 AA.  
AC ADB75186;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Prostate cancer marker protein.  
XX  
KW Prostate; cancer; cytostatic; gene therapy; marker.  
XX  
OS Homo sapiens.  
XX  
PN WO2003009814-A2.  
XX  
PD 06-FEB-2003.  
XX  
PF 25-JUL-2002; 2002WO-US023913.  
XX  
PR 25-JUL-2001; 2001US-0307982P.  
PR 22-AUG-2001; 2001US-0314356P.  
PR 25-SEP-2001; 2001US-0325020P.  
PR 12-DEC-2001; 2001US-0341746P.  
PR 05-MAR-2002; 2002US-0362158P.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;  
PI Hoersh S, Kamatkar S, Wonsey AW, Glatt K, Zhao X, Anderson D;  
XX  
XX WPI; 2003-248033/24.  
XX  
XX New nucleic acid molecule, useful for diagnosing or treating prostate  
XX cancer.  
XX  
XX Disclosure; SEQ ID NO 10; 99pp; English.  
XX  
XX The invention relates to newly discovered cancer markers associated with  
XX the cancerous state of prostate cells. Also disclosed is a method of  
XX assessing whether a patient is afflicted with prostate cancer. The method  
XX of the invention involves assessing whether a patient is afflicted with  
XX prostate cancer by comparing the level of expression of a marker in a  
XX patient sample and the normal level of expression of the marker in a  
XX control non-prostate cancer sample, where a significant increase in the  
XX level of expression of the marker in the patient sample and the normal  
XX level indicates that the patient is afflicted with prostate cancer.  
XX Nucleic acids of the invention are useful for diagnosing or treating  
XX prostate cancer, and may be useful in gene therapy. Sequences given in  
XX ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence  
XX data for this patent did not form part of the printed specification, but  
XX was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 734 AA;  
Query Match 100.0%; Score 3984; DB 7; Length 734;  
Best Local Similarity 100.0%; Pred. No. 5.5e-292;  
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MWVLFLLSGGLGRMDSPLPQVITVPEKIRSIIEKGIESQASVKIVIEGPKYTVNLM 60  
Db 1 MWVLFLLSGGLGRMDSPLPQVITVPEKIRSIIEKGIESQASVKIVIEGPKYTVNLM 60  
Qy 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGIYEGYPKSVVMVSTCTGLRGLQFE 120  
Db 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGIYEGYPKSVVMVSTCTGLRGLQFE 120  
Qy 121 NVSYGIEPLESSVGFPHYIYQVKHKADVSLYNEKDIERSDLSPKQASBPQDFAKYIE 180  
Db 121 NVSYGIEPLESSVGFPHYIYQVKHKADVSLYNEKDIERSDLSPKQASBPQDFAKYIE 180  
Qy 181 MHVIVEKOLYNHMGSDTTVAQVQFOLIGITNAIFVSFNITIIILSSLELWIDENKIATTG 240  
Db 181 MHVIVEKOLYNHMGSDTTVAQVQFOLIGITNAIFVSFNITIIILSSLELWIDENKIATTG 240  
Qy 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFOGKMCNDANYAGGVVLHPRT 300  
Db 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFOGKMCNDANYAGGVVLHPRT 300  
Qy 301 ISLESALVILAQLLSLMSGITYDDINKCQCSGAVCIMNPEAIHFSGVKIFSNCSFEDFAH 360  
Db 301 ISLESALVILAQLLSLMSGITYDDINKCQCSGAVCIMNPEAIHFSGVKIFSNCSFEDFAH 360  
Qy 361 FISKQKSOCLHNQPRLDPPFKQAVCGNAKLEAGEECDCGTEODCALIGETCCDIATCRF 420  
Db 361 FISKQKSOCLHNQPRLDPPFKQAVCGNAKLEAGEECDCGTEODCALIGETCCDIATCRF 420  
Qy 421 KAGSNCAEGPCCNCLFMSKERMCRPSFECDDLPEYCNSSASCPENHYVQTGHPCGLNQ 480  
Db 421 KAGSNCAEGPCCNCLFMSKERMCRPSFECDDLPEYCNSSASCPENHYVQTGHPCGLNQ 480  
Qy 481 WICIDGVCMSGDKQCTDTTFCGEVFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540  
Db 481 WICIDGVCMSGDKQCTDTTFCGEVFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540  
Qy 541 CGKLICKYVGKFLLOIPRATIIYANISGHLCIAVEFASDHADSQKMWIKDGTSCGSKNVC 600  
Db 541 CGKLICKYVGKFLLOIPRATIIYANISGHLCIAVEFASDHADSQKMWIKDGTSCGSKNVC 600  
Qy 601 RNORCVSSSYLGYDCTTDCNDRGVNCKKHCHCSASYLPPDCSVQSDLPWGSIDSGNF 660  
Db 601 RNORCVSSSYLGYDCTTDCNDRGVNCKKHCHCSASYLPPDCSVQSDLPWGSIDSGNF 660  
Qy 661 PPVAIPARLPERRYENIYHSPKMRWPFLLFIPFFIFCVLIAIMVKVNFQKKWRTEDY 720  
Db 661 PPVAIPARLPERRYENIYHSPKMRWPFLLFIPFFIFCVLIAIMVKVNFQKKWRTEDY 720  
Qy 721 SSDEQPESESEPKG 734  
Db 721 SSDEQPESESEPKG 734

RESULT 4  
ID ADB75184 standard; protein; 735 AA.  
XX  
AC ADB75184;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Prostate cancer marker protein.  
XX  
KW Prostate; cancer; cytostatic; gene therapy; marker.  
XX  
OS Homo sapiens.  
XX

```
PN WO2003009814-A2.
XX
PD 06-FEB-2003.
XX
PF 25-JUL-2002; 2002WO-US023913.
XX
PR 25-JUL-2001; 2001US-0307982P.
XX
PR 22-AUG-2001; 2001US-0314356P.
XX
PR 25-SEP-2001; 2001US-0325020P.
XX
PR 12-DEC-2001; 2001US-0341746P.
XX
PR 05-MAR-2002; 2002US-0362158P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
PI Hoersh S, Kamatkar S, Wonsey AM, Glat K, Zhao X, Anderson D;
XX
XX WPI; 2003-248033/24.
XX
XX New nucleic acid molecule, useful for diagnosing or treating prostate
PT cancer.
XX
XX Claim 4; SEQ ID NO 8; 99pp; English.
XX
XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.
CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 735 AA;
XX
Query Match 99.6%; Score 3969.5; DB 7; Length 735;
Best Local Similarity 99.7%; Pred. NO. 6.9e-291;
Matches 733; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 MW-VLFLSLGGLRMDNSFLPQITVPEKIRSIIEKIGIESQASKIVIEGKPYTVNL 59
DB 1 MWRVFLSLGGLGURMDNSFQSLPQITVPEKIRSIIEKIGIESQASKIVIEGKPYTVNL 60
QY 60 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGPKSVWVSTCTGLRGVLQF 119
DB 61 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGPKSVWVSTCTGLRGVLQF 120
QY 120 ENVSYGIEPLESSVGFEBHYIVQVRKKADVSLYNEKDIESRDLFSKLOSAEPQDFAKYI 179
DB 121 ENVSYGIEPLESSVGFEBHYIVQVRKKADVSLYNEKDIESRDLFSKLOSAEPQDFAKYI 180
QY 180 EMHVIVKQLYNHMGSDTTVAQKVFQVIGLITNAIFVSNITITLSLELWIDENKIATT 239
DB 181 EMHVIVKQLYNHMGSDTTVAQKVFQVIGLITNAIFVSNITITLSLELWIDENKIATT 240
QY 240 GEANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFQGMCDANVAGGVHLHPR 299
DB 241 GEANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFQGMCDANVAGGVHLHPR 300
QY 300 TISLESIAVLAQLLSLMGITYDDINKCQSGAVCMNPEAIHFGSKVIFSNCSFEDFA 359
DB 301 TISLESIAVLAQLLSLMGITYDDINKCQSGAVCMNPEAIHFGSKVIFSNCSFEDFA 360
QY 360 HFTSKQSKQCLHNPRLDPPFPKQAVCGNAKLEAGEBCDCGTEDCCALIGETCCDIATCR 419
DB 361 HFTSKQSKQCLHNPRLDPPFPKQAVCGNAKLEAGEBCDCGTEDCCALIGETCCDIATCR 420

QY 420 FKAGSNCAEGPCCECLFMSKRCRPSFECDLPEYCNSSASCSPENHYVQTGHPCLN 479
DB 421 FKAGSNCAEGPCCECLFMSKRCRPSFECDLPEYCNSSASCSPENHYVQTGHPCLN 480
QY 480 QWICIDGVCMSGDKQCTDTFTGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNL 539
DB 481 QWICIDGVCMSGDKQCTDTFTGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNL 540
QY 540 QCGKLICKYVGKFLQIPRATIIYANISGHLCTAVEFASDHADSQKMWIKDGTSCGSKNV 599
DB 541 QCGKLICKYVGKFLQIPRATIIYANISGHLCTAVEFASDHADSQKMWIKDGTSCGSKNV 600
QY 600 CRNORCVSSSYLGVDCTTDCNDRGVCNNKHKCHCSASYLPDCSVQSDLPWPGSIDSGN 659
DB 601 CRNORCVSSSYLGVDCTTDCNDRGVCNNKHKCHCSASYLPDCSVQSDLPWPGSIDSGN 660
QY 660 FPPVAIPARLPERRYIENIYHSPMRWPPFLFIPFFIIFCVLIAIMVKNVQFKKWRTE 719
DB 661 FPPVAIPARLPERRYIENIYHSPMRWPPFLFIPFFIIFCVLIAIMVKNVQFKKWRTE 720
QY 720 YSSDEQPESESEPKG 734
DB 721 YSSDEQPESESEPKG 735

RESULT 5
ADB75182
ID ADB75182 standard; protein; 753 AA.
XX
AC ADB75182;
XX
DT 04-DEC-2003 (first entry)
XX
DE Prostate cancer marker protein.
XX
KW Prostate; cancer; cytostatic; gene therapy; marker.
XX
OS Homo sapiens.
XX
XX WO2003009814-A2.
XX
PD 06-FEB-2003.
XX
PF 25-JUL-2002; 2002WO-US023913.
XX
PR 25-JUL-2001; 2001US-0307982P.
XX
PR 22-AUG-2001; 2001US-0314356P.
XX
PR 25-SEP-2001; 2001US-0325020P.
XX
PR 12-DEC-2001; 2001US-0341746P.
XX
PR 05-MAR-2002; 2002US-0362158P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
PI Hoersh S, Kamatkar S, Wonsey AM, Glat K, Zhao X, Anderson D;
XX
XX WPI; 2003-248033/24.
XX
XX New nucleic acid molecule, useful for diagnosing or treating prostate
PT cancer.
XX
XX Claim 4; SEQ ID NO 6; 99pp; English.
XX
XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.
CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 735 AA;
XX
Query Match 99.6%; Score 3969.5; DB 7; Length 735;
Best Local Similarity 99.7%; Pred. NO. 6.9e-291;
Matches 733; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 MW-VLFLSLGGLRMDNSFLPQITVPEKIRSIIEKIGIESQASKIVIEGKPYTVNL 59
DB 1 MWRVFLSLGGLGURMDNSFQSLPQITVPEKIRSIIEKIGIESQASKIVIEGKPYTVNL 60
QY 60 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGPKSVWVSTCTGLRGVLQF 119
DB 61 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGPKSVWVSTCTGLRGVLQF 120
QY 120 ENVSYGIEPLESSVGFEBHYIVQVRKKADVSLYNEKDIESRDLFSKLOSAEPQDFAKYI 179
DB 121 ENVSYGIEPLESSVGFEBHYIVQVRKKADVSLYNEKDIESRDLFSKLOSAEPQDFAKYI 180
QY 180 EMHVIVKQLYNHMGSDTTVAQKVFQVIGLITNAIFVSNITITLSLELWIDENKIATT 239
DB 181 EMHVIVKQLYNHMGSDTTVAQKVFQVIGLITNAIFVSNITITLSLELWIDENKIATT 240
QY 240 GEANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFQGMCDANVAGGVHLHPR 299
DB 241 GEANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFQGMCDANVAGGVHLHPR 300
QY 300 TISLESIAVLAQLLSLMGITYDDINKCQSGAVCMNPEAIHFGSKVIFSNCSFEDFA 359
DB 301 TISLESIAVLAQLLSLMGITYDDINKCQSGAVCMNPEAIHFGSKVIFSNCSFEDFA 360
QY 360 HFTSKQSKQCLHNPRLDPPFPKQAVCGNAKLEAGEBCDCGTEDCCALIGETCCDIATCR 419
DB 361 HFTSKQSKQCLHNPRLDPPFPKQAVCGNAKLEAGEBCDCGTEDCCALIGETCCDIATCR 420
```



```

Db 422 FGPSECYSHLNSKTDVSGNCGISGSGYTQCADNLQCGKLCCKVGRFLQIPRATIIYA 481
Qy 565 NISGHLCIAVEFADHADSQKMWIKDGTSCGSKVCRNQRQVSSSYLGVDCTTCKNDRG 624
Db 482 NISGHLCIAVEFADHADSQKMWIKDGTSCGSKVCRNQRQVSSSYLGVDCTTCKNDRG 541
Qy 625 VCNKKGHCHCSASYLPDPCSVQSDLPWPGGSDSGNFPFPAIPARLPERRYIENIYHSPKM 684
Db 542 VCNKKGHCHCSASYLPDPCSVQSDLPWPGGSDSGNFPFPAIPARLPERRYIENIYHSPKM 601
Qy 685 RWPFFLPFPIIFCVLIAIMVKNFQKWKRTEDYSSDEQPESESEPKG 734
Db 602 RWPFFLPFPIIFCVLIAIMVKNFQKWKRTEDYSSDEQPESESEPKG 651

RESULT 7
AAR87036
ID AAR87036 standard; protein; 735 AA.
XX
AC AAR87036;
DT 30-SEP-1996 (first entry)
XX
DE Mouse PH30 beta chain sperm protein.
XX
KW Mouse PH30 beta chain sperm protein; contraceptive;
KW QDE integrin binding domain.
XX
OS Mus musculus.
XX
FH Key
FT Region
FT 279..735
FT /note= "see AAR87035"
XX
EN W09535118-A1.
XX
PD 28-DEC-1995.
XX
PF 06-JUN-1995; 95WO-US007295.
XX
PR 20-JUN-1994; 94US-00264101.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Alves K, Gupta SK, Hollis GF;
XX
PT WPI; 1996-058212/06.
XX
DR N-PSDB; AAT07327.
XX
Human and mouse sperm protein PH30 beta chain and related DNA - useful in
contraceptive vaccines.
PS Example 2; Page 39-41; 85pp; English.
XX
CC Mouse PH30 beta chain sperm protein, having a QDE integrin binding
CC domain, is 55.2% identical to guinea pig PH30 beta. The protein may be
CC produced recombinantly and used in a contraceptive composition containing
CC an effective adjuvant and an amount of sperm protein which is effective
CC for the stimulation of antibodies which bind to sperm protein in vivo,
CC thereby preventing or substantially reducing the rate of sperm-egg fusion
XX
SQ Sequence 735 AA;

Query Match 61.2%; Score 2440; DB 2; Length 735;
Best Local Similarity 59.6%; Pred. No. 2.6e-175;
Matches 439; Conservative 119; Mismatches 167; Indels 12; Gaps 7;

Qy 3 VLFLLSG---LGGI---RMDSNFDSLPGVITVPEKIRSIKEGIESQASQKIVIEGKPYT 56
Db 4 ILLLSGLSELGGLSQSGTEGTRKHLVQVTVPEKIRSVTSNGYETQVYNLKEGTYT 63
Qy 57 VNLMOQNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYGQYIEGYPKSVVMVSTCTGLRGV 116

```

```

Db 64 LDLMQKFPDPNFRVSYDNAGIMRSLEQFQNICYFQGYIEGYPNSMWIVSTCTGLRGF 123
Qy 117 LQPNVSYGIEPLESSVGFVHVYQVGHKADVSLYNEKIDIESDLSEFKLSQAPQODFA 176
Db 124 LQFNVSYGIEPLESSVGFVHVYQVPEKGGALLYAEKIDIDLDLSQYKRSIKPQIRVS 183
Qy 177 KYIEMHVIVKQLNYNHMGSDTTVVAAQVFLQIGLTNAIFVSFNITITILSSLELWIDENKI 236
Db 184 HYLEIHIVVEKQFHEHIGADTAIVTQKIFQIGLANAIFAPFNLTVILSSLEFWMDENKI 243
Qy 237 ATTCEANELLHTFLRWKTSYLIVLRPHDVAFLLYVREKSNVYVATFQGMCDANVAGGVVL 296
Db 244 LTTGDANKLJYRFLKWKQSYLVLRPHDMAFLLYVRYNTTDYVATYQGMCDKNYAGGVAL 303
Qy 297 HPTISLESIAVILLAQLLSMGITDDINKCCSGAVCIMNPEAIHFSGVKIFSNCSFE 356
Db 304 HPKAVTLESIAVILVQLLSMGLAYDDVNNKCCGVPCVMNPEAPHSSGVRAFSNCSME 363
Qy 357 DFAHFISKQKSQLHNQPRLDFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIA 416
Db 364 DFSKFITSQSSHCLQNOPTLQPSYK-MAVCGNGEVEBEICDQG-KKGCAEMPPPPCCNPD 421
Qy 417 TCRPKAGSNCAEGPCCECLFMSKERMCRPSFECDLPEYCNCGSSASCENHYVOTGHP 476
Db 422 TKLSLDGSESGGICCNCKLRKGEVCRLAODECDVTEYCNGTSEVC-BDFFVQNGHPC 480
Qy 477 GLNQWICIDGVCMGDKQCTDTFGKEVFGPSECYSHLNSKTDVSGNCGISDSYTCQEA 536
Db 481 DNRKWICINTCOSGEQQCDLFGIDAGFSSECFWELNSKSDISGSCGISAGGYKCEPP 540
Qy 537 DNLQCGKLI CKYVGKFLQIPRATIIYANISGHLCIAVEFADHADSQKMWIKDGTSCGS 596
Db 541 NDRMCGKII CKYQSENILKRSATVIYANISGHVCSLEYVPGQHNSQKMWVRDGTVCGS 600
Qy 597 NKVCRNQRVSSSYLGVDCTTCKNDRGVGNKHKHCHCSASYLPDPCSVQSDLPWPGSID 656
Db 601 NKVCQKQKCVADTFLGYDCNLEKCNHHGVCNKKNCNCHCDPTLPPDCRKRKDSYPGSID 660
Qy 657 SGNFPPVAIPARLPERRYIENIYHSPKMRWPFPIFPIFFIIFCVLIAIMVKNFQKWK 716
Db 661 SGN-KERAEP--IPVRPIASRYRSKSPRWPFFLIIPFYVILVILGLVKNVYSQRMKWR 717
Qy 717 TEDYSDEQPESESEPK 733
Db 718 MDDFSSEQFESESESK 734

RESULT 8
AAU83636
ID AAU83636 standard; protein; 787 AA.
XX
AC AAU83636;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human PRO protein, Seq ID No 90.
XX
KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
KW breast cancer; prostate tumour; rectal tumour; liver tumour;
KW pericyte cell proliferation; chondrocyte cell proliferation;
KW tumour necrosis factor-alpha.
XX
OS Homo sapiens.
XX
EN W0200208288-A2.
XX
PD 31-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-US021066.
XX
PR 20-JUL-2000; 2000US-0219556P.
PR 25-JUL-2000; 2000US-0220585P.

```







[illegible]

XX	Human; secreted and transmembrane protein; PRO; cardiac; cytosolic; antiangiogenic; hypotensive; vulnery; antiarteriosclerotic; gene therapy; cardiovascular disorder; endothelial disorder; angiogenic disorder; cardiac hypertrophy; trauma; cancer; age-related macular degeneration; atherosclerosis; hypertension; arterial restenosis; rheumatoid arthritis; angina; myocardial infarction; thrombophlebitis; lymphangitis; tumour angiogenesis; breast carcinoma; liver carcinoma; wound healing; chromosome mapping; gene mapping.
XX	Homo sapiens.
OS	US2003088063-A1.
PN	08-MAY-2003.
XX	12-AUG-2002; 2002US-00019003.
PF	25-JUL-2000; 2000US-0220664P.
XX	01-JUN-2001; 2001WO-US017800.
PR	29-JUN-2001; 2001WO-US021066.
PR	09-APR-2002; 2002US-00119480.
XX	(GETH ) GENENTECH INC.
PA	Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI, WPI, 2003-393229/37.
XX	N-PSDB; ACA68541.
DR	One hundred and eighty seven nucleic acids encoding PRO polypeptides, useful in diagnosis and treatment of cardiovascular (e.g. myocardial infarction), endothelial or angiogenic disorders in a mammal.
PT	Claim 11; Fig 90; 314pp; English.
XX	The invention describes one hundred and eighty seven nucleic acids encoding novel human secreted and transmembrane (PRO) polypeptides. The PRO nucleic acids, polypeptides, agonists and antagonists are useful for treating or diagnosing a cardiovascular, endothelial or angiogenic disorder in a mammal, e.g. cardiac hypertrophy, trauma, cancer, age-related macular degeneration, atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis, angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour angiogenesis (such as breast carcinoma and liver carcinoma) and wound healing. The PRO nucleic acids have applications in molecular biology, including use as hybridisation probes, and in chromosome and gene mapping. This is the amino acid sequence of a novel human secreted and transmembrane PRO polypeptide
XX	Sequence 787 AA;
SQ	Query Match 38.3%; Score 1524.5; DB 6; Length 787; Best Local Similarity 41.1%; Pred. No. 4.3e-106; Matches 310; Conservative 134; Mismatches 259; Indels 51; Gaps 18
QY	1 MWVFLFLSGGLRMD-SNPDLSLVQITVPKIRSIIEGIE---SQASKIVIEGKPYT 56
Db	4 LWL--LIAGLCGLLASRPGFQNSLLQIVPKIKTNTNDSSEIEYEQISYIIPDEKUYT 61
QY	57 VNLNQKFLPHNFRVSYSGTGMKPLDQDFNQFCHYQGYEGYPKSVVMVSTCTGLRGV 116
Db	62 VHLKQRYFLADNFMVILYN-QGSMTYSSDIQTQCYQGNIEGYPDSMTVLSTCSGLRGI 120
QY	117 LQFNVSYGTEPLESSVGFEHVIYQVGHKKADVLYNEKDIETR---DLSFKLSAEPQQ 173
Db	121 LQFNVSYGTEPLESSAVEFQVLYKLNEDNDNDIAIFDRSLKEQPMDDNIFISEKSEPAV 180
QY	174 D--FAKYIEHVIYKOLYNHMGSDTTVAQKVFQILGTLTAIFVSFNITILSLELWI 231
Db	181 PDLEPLFLEHHIVVDKTLTIFWGSDSMIVTNKVIETVGLANSMTQFKVTIVLSLELWS 240
QY	232 DENKIATTGEANELLHTFLRWKTSYLVLRPHDVAFLVLYREKSNVVGATFGKMKDANYA 291



```
RESULT 13
ABJ72400
ID ABJ72400 standard; protein; 787 AA.
XX
AC ABJ72400;
XX
XX
DT 06-NOV-2003 (first entry)
XX
DE Human PRO21340 protein.
XX
KW PRO; blood; proliferation; pericyte cell; TNF alpha; chondrocyte;
KW tumour necrosis factor; proliferation; differentiation; gene therapy;
KW dermal fibroblast.
XX
XX
OS Homo sapiens.
XX
PN US2003027988-A1.
XX
PD 06-FEB-2003.
XX
PF 26-AUG-2002; 2002US-00227884.
XX
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
DR N-PSDB; ABT44553.
XX
DR WPI; 2003-503301/47.
XX
XX
PT New PRO protein encoding nucleic acid, useful for preparing PRO
PT polypeptides and anti-PRO antibodies for detecting the presence of a
PT tumor in a mammal.
XX
PS Claim 11; Fig 90; 324pp; English.
XX
CC The invention relates to a novel isolated PRO protein encoding nucleic
CC acid. The nucleic acid of the invention may be useful for preparing PRO
CC polypeptides and anti-PRO antibodies for detecting the presence of a
CC tumour in a mammal. Furthermore, the molecules of the invention may be
CC useful for stimulating proliferation or gene expression in pericyte
CC cells, the release of tumour necrosis factor (TNF)-alpha from human
CC blood, the proliferation or differentiation of chondrocyte cells and for
CC inhibiting the proliferation of normal human dermal fibroblast cells.
CC Finally, the molecules may be utilised during gene therapy. The current
CC sequence is that of the human PRO protein of the invention
XX
SQ Sequence 787 AA;

Query Match 38.3%; Score 1524.5; DB 6; Length 787;
Best Local Similarity 41.1%; Pred. No. 4.3e-106;
Matches 310; Conservative 134; Mismatches 259; Indels 51; Gaps 18;

QY 1 MWVLFLLSLGGLGRLMD-SNFDLPVQITVPEKIRSIKEGIE---SQASVKIVIEGKPYT 56
DB 4 LML--LLAGLCGLARPGQNSLLQIVPEIKITNTDSSEIEYQISYIIPIDELKLT 61
QY 57 VNLQMKNFLPHNFRVYSYSGTGMKPLDQDFQNFCHYQGYIEGPKSVVMVSTCTGLRGV 116
DB 62 VHLKQRYFLADNPWILYN-QGSMNTYSSDIQTCYQYQGNIEGYPDSMTLSTCSGLGI 120
QY 117 LQENVSYGIEPLESSVGFEHYIQVKKHKAADVSLYNEKDIESR---DLSEFKLQSAEPOQ 173
DB 121 LQENVSYGIEPLESAVEFOHLYLKLKNEKDIAIFIDRLSKLEQPMDDNIFISEKSEPAV 180
QY 174 D--FAKYEMHIVAEKQLYNHMGSDTIVVAQKVFQILGLTNAIFVFNITIISSLELWI 231
DB 181 PDLFPLYLEMHIVVDKTLVDYWGSDSMIVTNKVIIEIVGLANSMTQFKVTIVLSLELWS 240
```

---

```
QY 232 DENKIATTGEANELLHTFLRWKTSYLVLRPHDVAFLLVYREKSNYVGATPQGMKCDANYA 291
DB 241 DENKISTVGEADELLQKFLWKQSYLNLAPHDIAYLLIYMDYPRYLGAVPFGTMCITRYS 300
QY 292 GGVVLHPRTISLESIAVLAQLLSMGITYDDINKCOCSGAVCINWPEAHFSGVKIFS 351
DB 301 AGVALYPKEITILEAFVITQMLAUSLGISYDDPKKQCQSESTCINWPEVVGNSGKVTFS 360
QY 352 NCSFEDFAHFISKQKSQLHNQPRLPDPFFKQAVCGNAKLEAGEECDCGTEQDCALIGET 411
DB 361 SCSLRSFQNFISNVGVKCLQNKRPQMK-KSPKPVCGNGRLEGNEICDCGTEAQCQ--PAS 417
QY 412 CDDIATCRFKAGSNCAEGPCCNCLFMSKERMCRP-SFEECDLPYVCGNSGSSASCENHYV 470
DB 418 CDDFRTCVLKDGAKCYKGLCKCKQCLQSGVECRPKAHPEDCIAENCNGSSSPCCGPDITL 477
QY 471 QTGHPCGLNQWICIDGVCMGDKQCTDTFGKEVFGPSECYSHLNSKTDVSGNCGIS-DS 529
DB 478 INGLCKNNKPFICYDGDCHDLDAECESVFGKSRNAPFACYEEIQSQDRFGNCGDRDN 537
QY 530 GYTQCEADNLQCGKLICTKVGVKFLQIPRATIIYANISGHLCTIAVBFASDHADSQMKWIK 589
DB 538 KYVFCGWRNLI CGLVCTYPTKPFHQENGDIYAFVDSVCITVDYKLPRTVPDPLAVK 597
QY 590 DGTSCGSKNVCNRQRCVSSYL---GYDCTTDKNDRGVNNKHKHCHCSASYLPDPCSVQ 646
DB 598 NGSQCDIGRVCVNRCEVESRIIKASAHVC-SQCSGHGVCDNRKNCCHCSGPGYKPPNCQIR 656
QY 647 SLMWPGGSI---DSGNFPPVAIPARLPERRVNIENIYHSPMRW--PFELFIFFFIFCV 700
DB 657 S---KGFSIPPEEDMGSI-----MERASGKTENTWLLGLFLIALPILIV--- 696
QY 701 LIAIMVKNVFORKKWRTEDYSDEQPESESEPKG 734
DB 697 -TTAIVLARKQLKWFAKE---EEFPSSSEKSEG 726

RESULT 14
ABO34295
ID ABO34295 standard; protein; 787 AA.
XX
AC ABO34295;
XX
DT 19-SEP-2003 (first entry)
XX
DE Human secreted/transmembrane polypeptide PRO 21340.
XX
KW Human; chondrocyte stimulation; TNF-alpha stimulation; gene therapy;
KW human dermal fibroblast stimulation; tumour; tissue typing;
KW affinity purification.
XX
OS Homo sapiens.
XX
XX
FN US2003044934-A1.
XX
PD 06-MAR-2003.
XX
PF 28-AUG-2002; 2002US-00230338.
XX
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
XX
DR WPI; 2003-492274/46.
DR N-PSDB; ACD82220.
XX
PT New transmembrane polypeptides and nucleic acids encoding the
```

PT polypeptides, useful in gene therapy, in chromosome identification, as  
PT chromosome markers, or in generating probes.  
XX Claim 19; Fig 90; 315pp; English.  
XX The invention relates to an isolated nucleic acid encoding a PRO  
CC polypeptide. Nucleic acids that encode PRO can be used to generate either  
CC transgenic animals or knock-out animals useful in developing and  
CC screening of therapeutically useful reagents. The nucleic acids may also  
CC be used in gene therapy for replacing defective gene, in chromosome  
CC identification, as chromosome markers, or in generating probes to isolate  
CC full length PRO cDNA. The PRO polypeptides are useful for chondrocyte  
CC stimulation, TNF-alpha stimulation, human dermal fibroblasts stimulation  
CC and for detecting the presence of tumour in an animal. The PRO  
CC polypeptides are useful as molecular markers for protein electrophoresis  
CC and the isolated nucleic acids may be used for recombinantly expressing  
CC those markers. The PRO polypeptides and nucleic acids may also be used in  
CC tissue typing. Anti-PRO antibodies are useful in diagnostic assays for  
CC PRO and in affinity purification of PRO from recombinant cell culture or  
CC natural sources. The present sequence represents the amino acid sequence  
CC of a human secreted/transmembrane PRO polypeptide  
XX Sequence 787 AA;  
SQ  
Query Match 38.3%; Score 1524.5; DB 6; Length 787;  
Best Local Similarity 41.1%; Pred. No. 4.3e-106;  
Matches 310; Conservative 134; Mismatches 259; Indels 51; Gaps 18;  
QY 1 MWVFLSLGSGGLRMD-SNFDSLVPQITVPEKIRSIKEGIE---SQASKYIVIEGKPYT 56  
DB 4 LWL--LLAGLCGLASRPGFQNSLLQIVPEKIQTNTDSIEYEIQISYIIPIDEKLYT 61  
QY 57 VNLMOKNFLPHNRVSYSGTGIMKPLDQDFQFCHQYIEGVKSVVMVSTCTGLGV 116  
DB 62 VHLKQRYFLADNFMVLYLN-QGSNMTYSDDIQYIYQGNIEGYDPSMVTLSLCSGLGI 120  
QY 117 LQFENSYGIEPLESSVGFPHYIVQVHKKADVSLYNEKIDIESR--DLSEFKLQSAEPQ 173  
DB 121 LQFENSYGIEPLESAVEFQVLYLKLKEDNDIAIFIDRSLSKQPMDDNIFISEKSEPAV 180  
QY 174 D--FAKYIEMHIVIEKOLYNHMSDDTTVAQKVPQLIGLTNAIPVSNITIISSLELWI 231  
DB 181 PDLFPVLEHIVVDKTLVDYWGSDSMIVTNKVIIEIVGLANSMTQPKVIVLSSLELWS 240  
QY 232 DENKIATGTGANELLHTFLWKTSLVLRHDVLAFLAVREKSNYVCATQFGKMCADNYA 291  
DB 241 DENKISTVGEADELLQFLWKQSLNLRPHDIAIYLLIYMDYPRYLGAVPFGMTITRYS 300  
QY 292 GGVVLPRTISLESIAVLAQLSLSMGITVDDINKQCQSGAVCIMNPEAIHFSGVKIFS 351  
DB 301 AGVALYPEITLEAFVIVTQMLALSLSIGSYDDPKKQCSESTCIMNPEVQVNGVKTFS 360  
QY 352 NCSFEDFAHFISKQKSLNQLNPLRDPFFQKQAVCGNAKLEAGECDGTEQDCALIGET 411  
DB 361 SCSLRSQNFISNVGVKCLQNKPMQK-KSPKPVCGNGRLEGNIEICDCEGTAQCG--PAS 417  
QY 412 CCDIATCRFAGNSCARGPCENCLFMSKERMCP-SFESCDLPEYCNCGSSASCENHYV 470  
DB 418 CDFPRTVLDGAKCYKGLCCQCOILQSGVECHPKAHPEDIAENCGSPCEGPDITL 477  
QY 471 QTGHPGCLNWCICDGVCMGSKQCTDTTFGEVFGFPGSECVSHLNSKTDVSGNGIGS-DS 529  
DB 478 INGLSKNNKFCYDGDCHDLARCESVFGKSRNAPFACVEEYEQSQSDRFGNGCRDRN 537  
QY 530 GYTQCEADNLCQGLIKYGVKFLQIIPRATIIIVANISGHLCTIAVEFASHADSKRWIK 589  
DB 538 KYVFCGRNLTICGLVCTVTPRKPFHQENGVDVYAFVRDSVCITVDYKLPRTVPDPLAVK 597  
QY 590 DGTSCGSKNVCNRCQVSSVSL---GYDCTTDKNDRGVGNCKKHCHCSAYLPPDCSVQ 646  
DB 598 NGSCQDITGRVCNRECVSRITIKASAVC-SQQCSGHGVCDSRNKCHCSGPCYKPNQIR 656  
QY 647 SLDLPFGGSI-----DSGNFPFVPAIPARLPERRYENIYHSPMRW--PFFLPFPFIIFCV 700

DB 657 S---KGFSIFPEEDMGSI-----MERASGKXTNTWLLGFLIALPILIV--- 696  
QY 701 LTAIVKVNFORFKWRTEDYSSDEQSESESEPKG 734  
DB 697 -TTAIVLARKQDKWFAKE---EEFSSSEKSEG 726  
RESULT 15  
ABJ72102  
ID ABJ72102 standard; protein; 787 AA.  
XX AC ABJ72102;  
XX DT 16-OCT-2003 (first entry)  
XX DE Human membrane bound receptor/protein PRO211340 amino acid sequence.  
XX KW Human; PRO; membrane bound protein; membrane bound receptor;  
KW cell proliferation; cell migration; cell differentiation;  
KW mitogenic factor; survival factor; cytotoxic factor;  
KW differentiation factor; neuroepithelial; hormone; cell receptor;  
KW receptor-ligand interaction; cytostatic; chondrocyte; tumour.  
XX OS Homo sapiens.  
XX PN US2003065147-A1.  
XX PD 03-APR-2003.  
XX PF 29-AUG-2002; 2002US-00232224.  
XX PR 28-JUL-1999; 99US-0146222P.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-APR-2002; 2002US-00119480.  
XX (GETH ) GENENTECH INC.  
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;  
XX WPI: 2003-522018/49.  
XX N-PSDB; AB143926.  
PT One hundred and twenty two nucleic acids encoding PRO polypeptides,  
PT useful for the manufacture of a medicament for diagnosing or treating  
PT tumor.  
XX Claim 11; Fig 90; 315pp; English.  
XX This invention relates to one hundred and twenty two novel nucleic acids  
CC encoding human PRO membrane bound proteins or receptors. Extracellular  
CC proteins play important roles in the formation, differentiation and  
CC maintenance of multicellular organisms. The fate of many individual cells  
CC (for example proliferation, migration or differentiation) is typically  
CC governed by information received from other cells and the immediate  
CC environment. The information is often transmitted by secreted  
CC polypeptides (for example mitogenic factors, survival factors, cytotoxic  
CC factors, differentiation factors, neuroepithelial and hormones) which are  
CC received and interpreted by diverse cell receptors or membrane bound  
CC proteins. These membrane bound proteins and receptors may be of use as  
CC pharmaceutical and diagnostic agents, such as in the blocking of receptor  
CC -ligand interactions. The current invention provides the amino acid  
CC sequences of novel human membrane bound receptors and proteins, along  
CC with the cDNA sequences encoding them. The novel proteins of the  
CC invention may have cytostatic activities through the stimulation of  
CC chondrocytes. The nucleic acids of the invention may be useful for the  
CC manufacture of a medicament for diagnosing or treating a tumour in a  
CC mammal. In addition, they may be useful for measuring or detecting the  
CC expression of a tumour associated gene. The present sequence is the amino

Search completed: January 10, 2005, 22:22:24  
Job time : 163 secs

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:07:16 ; Search time 43 Seconds  
(without alignments)  
1642.398 Million cell updates/sec

Title: US-10-054-683-19  
Perfect score: 3984  
Sequence: 1 MWVFLSLGLGLRMDNF.....WRTDYSSDQPESESEPKG 734  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3984	100.0	734	2	JC4861
2	3687.5	92.6	735	2	G02937
3	1383.5	34.7	756	2	S47656
4	1365.5	34.3	736	2	S47645
5	1292	32.4	823	2	S18968
6	1288	32.3	777	2	I48100
7	1166	29.3	357	2	S23403
8	1097	27.5	655	2	JC7850
9	1031.5	25.9	825	2	S55060
10	1022.5	25.7	905	2	S55059
11	1016.5	25.5	660	2	S71949
12	857	21.5	903	2	S60257
13	795	20.0	600	2	I49281
14	766	19.2	826	2	A60385
15	762	19.1	609	2	S55270
16	729	18.3	617	2	S48160
17	724.5	18.2	732	2	I52361
18	721	18.1	814	2	G02930
19	719.5	18.1	571	2	S24789
20	711	17.8	735	2	I48101
21	710.5	17.8	616	2	A55796
22	698	17.5	776	2	S28258
23	696.5	17.5	789	2	S28259
24	695	17.4	670	2	I65967
25	692.5	17.4	713	2	I65253
26	663	16.6	610	2	JC7530
27	653.5	16.4	610	2	JC8056
28	641	16.1	952	2	T18900
29	616.5	15.5	429	2	A42972

30	614.5	15.4	416	2	A37877
31	606	15.2	419	2	A41607
32	594.5	14.9	419	2	A59414
33	584	14.7	549	2	S48169
34	568	14.3	473	2	I49283
35	559.5	14.0	524	2	S38539
36	553	13.9	478	2	A43296
37	534	13.4	484	2	JC8020
38	533	13.4	480	1	A30065
39	530.5	13.3	481	2	JC4342
40	529.5	13.3	481	2	S43125
41	501	12.6	478	2	JQ1301
42	486.5	12.2	1042	2	T26844
43	467	11.7	478	2	JC4880
44	393.5	9.9	414	2	S41608
45	392.5	9.9	411	1	HYSNFA

ALIGNMENTS

RESULT 1

JC4861  
fertilin beta chain - human  
C:Species: Homo sapiens (man)  
C:Date: 15-Aug-1996 #sequence\_revision 18-Oct-1996 #text\_change 09-Jul-2004  
C:Accession: JC4861  
R:Gupta, S.K.; Alves, K.; O'Neil Palladino, L.; Mark, G.E.; Hollis, G.F.  
Biochem. Biophys. Res. Commun. 224, 318-326, 1996  
A:Title: Molecular cloning of the human fertilin beta subunit.  
A:Reference number: JC4861; MUID:96295488; PMID:8702389  
A:Accession: JC4861  
A:Molecule type: mRNA  
A:Residues: 1-734 <GUP>  
A:Cross-references: UNIPROT:Q99965; GB:U38805; NID:G4151118; PIDN:AAD04206.1; PID:G41511.1  
C:Comment: This protein is an integral sperm membrane glycoprotein, and plays a role in  
C:Superfamily: mouse meltrin alpha; disintegrin homology  
C:Keywords: Glycoprotein; integrin binding; transmembrane protein  
F:382-734/Product: fertilin beta chain #status predicted <NAT>  
F:382-467/Domain: disintegrin homology <DIS>  
F:448-450/Region: integrin binding #status predicted  
F:686-708/Domain: transmembrane #status predicted <TM>  
F:121,219,352,458,565/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	100.0%;	Score 3984;	DB 2;	Length 734;
Best Local Similarity	100.0%;	Pred. No. 5.3e-265;		
Matches	734;	Conservative	0;	Mismatches 0;
				Indels 0; Gaps 0;
Qy	1	MWVFLSLGLGLRMDNF	SLPVQITVPEKIRSI	IIKEGIESQASYKIVIEGKPYTNLM 60
Db	1	MWVFLSLGLGLRMDNF	SLPVQITVPEKIRSI	IIKEGIESQASYKIVIEGKPYTNLM 60
Qy	61	QKQFLPHNFRVYS	SGTGIMKPLDDQDF	QNFCHQYIEGYPKSVVMYSTCTGLRGVLQFE 120
Db	61	QKQFLPHNFRVYS	SGTGIMKPLDDQDF	QNFCHQYIEGYPKSVVMYSTCTGLRGVLQFE 120
Qy	121	NVSGIEPLESSVGF	EHVIYQVKHKADV	SLYNEKDIIESRDLSPKLSAQSPQDFAKYIE 180
Db	121	NVSGIEPLESSVGF	EHVIYQVKHKADV	SLYNEKDIIESRDLSPKLSAQSPQDFAKYIE 180
Qy	181	MHVIVEKQLNHN	GSDDTTVAQKVF	QLIGLTNAIFVFNITII
Db	181	MHVIVEKQLNHN	GSDDTTVAQKVF	QLIGLTNAIFVFNITII
Qy	241	EANELHTFRLW	KTSYLVLRPHD	VAFLLVYREKSNYVGATFQGMCDANTAGGVVLHPRT 300
Db	241	EANELHTFRLW	KTSYLVLRPHD	VAFLLVYREKSNYVGATFQGMCDANTAGGVVLHPRT 300
Qy	301	ISLESVILAQ	LLSLSMGITYDD	INKCQCSGAVCINNP
Db	301	ISLESVILAQ	LLSLSMGITYDD	INKCQCSGAVCINNP
Qy	361	FISKQKSC	LIHNPRLDP	PFKQAVCGNAKLEAGECDGTEQDCALIGETCCDIATCRF 420
Db	361	FISKQKSC	LIHNPRLDP	PFKQAVCGNAKLEAGECDGTEQDCALIGETCCDIATCRF 420

```
Db 361 FISKQKSQLHNQRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420
Qy 421 KAGSNCAEGCCENCLFMSKERMCRPSFECDDLPYCNCGSSASCENHYVQTGHPGCLNQ 480
Db 421 KAGSNCAEGCCENCLFMSKERMCRPSFECDDLPYCNCGSSASCENHYVQTGHPGCLNQ 480
Qy 481 WICIDGCMGSDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNGIGSDSGVTQCEADNLQ 540
Db 481 WICIDGCMGSDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNGIGSDSGVTQCEADNLQ 540
Qy 541 CGKLCIKYVGKFLQIPRATIIYANISGHLCAVEFASDHADSKQMWIKDGTSCGSKNVC 600
Db 541 CGKLCIKYVGKFLQIPRATIIYANISGHLCAVEFASDHADSKQMWIKDGTSCGSKNVC 600
Qy 601 RNQRCVSSSYLGVDCTTDKNDRGVNNKXKHCCHCSASYLPPDCSVQSDLPFGGSDSGNF 660
Db 601 RNQRCVSSSYLGVDCTTDKNDRGVNNKXKHCCHCSASYLPPDCSVQSDLPFGGSDSGNF 660
Qy 661 PPVAIPARLPERRYIENIYHSPKMRWPFELFIPFFIIFCVLIATIMVKVNFQKRWRTEDY 720
Db 661 PPVAIPARLPERRYIENIYHSPKMRWPFELFIPFFIIFCVLIATIMVKVNFQKRWRTEDY 720
Qy 721 SSDEQPESESEPKG 734
Db 721 SSDEQPESESEPKG 734

RESULT 2
G02937
fertilin beta - crab-eating macaque
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
R;Perry, A.C.F.; Gichuhi, P.M.; Jones, R.; Hall, L.
R;Ramaraio, C.S.; Myles, D.G.; White, J.M.; Primakoff, P.
submitted to the EMBL Data Library, August 1995
A;Reference number: G12615
A;Accession: G02937
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-735 <RAM>
A;Cross-references: UNIPROT:Q28478; EMBL:U39959; NID:G998339; PID:G998340
Biochem. J. 307, 843-850, 1995
A;Title: Cloning and analysis of monkey fertilin reveals novel alpha subunit isoforms.
A;Reference number: S55059; MUID:95260313; PMID:7741716
A;Accession: S55061
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-722, 'S', 724-735 <PER>
A;Cross-references: EMBL:X77653; NID:G794076; PIDN:CAA54733.1; PID:G794077
C;Superfamily: mouse meltrin alpha; disintegrin homology
F;383-468/Domain: disintegrin homology <DIS>

Query Match 92.6%; Score 3687.5; DB 2; Length 735;
Best Local Similarity 91.3%; Pred. No. 1.1e-244;
Matches 671; Conservative 31; Mismatches 32; Indels 1; Gaps 1;

Qy 1 MW-VLFLSLGLGLRMDNSFSLPVQITVPKIRSIKEGIESQASQKIVIEGKPYTNVL 59
Db 1 MWRVFLSLGLGLRMDNSFSLPVQITVPKIRSIKEEIESQVSQKIVIEGKPYTNAL 60
Qy 60 MOKNPLPHNFVSYSGTGIMKPLDQDFQNFCHQYIEGPKSVVMVSTCTGLRGVLQF 119
Db 60 MOKNPLPHNFVSYNGTGMKPLDQDFQNFCHQYIEGPKSVVMVSTCTGLRGVLQF 120
Qy 120 ENVSYGIEPLESSVGFHEVIVQVKKADADVLYNEKDIERSDLSPKLSAQSPQDFAKYI 179
Db 121 ENVSYGIEPLESSVGFHEVIVQVKKADADVLYNEKDIERSDLSPKLSQIEPQDFAKYI 180
Qy 180 EMHVIVEKQLYNHMGSDTTVAQKVPQLIGLTNAIFVSNITILSSLELWIDENKIATT 239
Db 181 EMHVIVEKQLYNHMGSGTTVTQKIFQLIGLTNAIFVSNITILSSLELWIDENKIATT 240
```

```
Qy 240 GBANELLLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFGKMKCDANYAGGVVLHPR 299
Db 241 GDAKELLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFGKMKCDANYAGGVVLHPR 300
Qy 300 TISLESIAVILAQLLSLSMGIYDDINKQCCSGAVCIMNPEAIHFSGVKIFNSCSEDEFA 359
Db 301 TISLESIAVILAQLLSLSMGIYDDINKQCCSGAAVCIMNPEAIHFSGVKIFNSCSEDEFA 360
Qy 360 HFTSKQKSOCLHNQRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCR 419
Db 361 HFTSKQKSOCLHNQRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCR 420
Qy 420 FRAGSNCAEGPCENCLFMSKERMCRPSFECDDLPYCNCGSSASCENHYVQTGHPGCLN 479
Db 421 FRAGSNCAEGPCENCLFMSKERMCRPSFECDDLPYCNCGSSASCENHYVQTGHPGCLN 480
Qy 480 QWICIDGCMGSDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNGIGSDSGVTQCEADNL 539
Db 481 QWICIDGCMGSDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNGIGSDSGVTQCEADNL 540
Qy 540 QCGKLCIKYVGKFLQIPRATIIYANISGHLCAVEFASDHADSKQMWIKDGTSCGSKNVC 599
Db 541 QCGKLCIKYVGKFLQIPRATIIYANISGHLCAVEFASDHADSKQMWIKDGTSCGSKNVC 600
Qy 600 CRNQCVRSSSYLGVDCTTDKNDRGVNNKXKHCCHCSASYLPPDCSVQSDLPFGGSDSGN 659
Db 601 CRNQCVRSSSYLGVDCTTDKNDRGVNNKXKHCCHCSASYLPPDCSVQSDLPFGGSDSGN 660
Qy 660 FPPVAIPARLPERRYIENIYHSPKMRWPFELFIPFFIIFCVLIATIMVKVNFQKRWRTED 719
Db 661 FPLVAVPARLPERRHENVYHSPKMRWPLFLFIPFFIIFCVLIATIMVKVNFQKRWRTED 720
Qy 720 YSSDEQPESESEPKG 734
Db 721 YSTDEQPESESEPKG 735

RESULT 3
S47656
tMDC II protein - crab-eating macaque
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S47656
R;Perry, A.C.F.; Barker, H.L.; Jones, R.; Hall, L.
Biochim. Biophys. Acta 1207, 134-137, 1994
A;Title: Genetic evidence for an additional member of the metalloproteinase-like, disint
A;Reference number: S47656; MUID:94318664; PMID:8043604
A;Accession: S47656
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-756 <PER>
A;Cross-references: UNIPROT:Q28483; EMBL:X77619; NID:G531477; PIDN:CAA54713.1; PID:G53147
C;Superfamily: mouse meltrin alpha; disintegrin homology
F;388-473/Domain: disintegrin homology <DIS>

Query Match 34.7%; Score 1383.5; DB 2; Length 756;
Best Local Similarity 38.1%; Pred. No. 6.6e-87;
Matches 293; Conservative 122; Mismatches 290; Indels 65; Gaps 18;

Qy 3 VLFLSLGLGLRMDNSFSLPVQITVPKIRSI-IEGIESQASQKIVIEGKPYTNLMQ 61
Db 4 LLVLLTGLGGMHADLNPHTFTLTQTIPEKISSDAKTDPEHNVVYMITIEGKPYFVHLKK 63
Qy 62 KNFLPHNFVSYSGTGI--MKPLDQDFQNFCHQYIEGPKSVVMVSTCTGLRGVLQF 119
Db 64 QSLSSASFTHSDKMDIRHSPF--LVQMDCNVNGYVAGIPNSLTVLSCVSGLRGTML 121
Qy 120 ENVSYGIEPLESSVGFHEVIVQVKKADADVLYNEKDIERSDLSPKLSAQSPQDFAK-- 177
Db 122 KNTSYGIEPLEMEASVGFTHKIYEKFDATNILLEENDYSWFNSEYQVRKSEKTDFTKLF 181
Qy 178 --YIEMHVIVEKQLYNHMGSDTTVAQKVPQLIGLTNAIFVSNITILSSLELWIDENK 235
```

```
Db 182 PRYEMHVVVDKDLFDYMGSDINAVTKVQIIGLVNTMLTQLQTLTVIISIEIWSNKNK 241
Qy 236 IATTGEANELHFTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFOGKMCNDANYAGGVV 295
Db 242 ISTTGHAEYVLLBEFFEWKDHLPKPHQIAYLFFYRKLPFLIGATFFGQVCKNDKFAAVA 301
Qy 296 LHPRTISLESIAVLAQLLSLMSGITYDDINKCOCGAVCIMPEAHFSGVKIFNSCSF 355
Db 302 LYPEGLSLESYTVIIVQLLGLNLGLTDTCHCSGDVCTMTPKAVSYGSGVDFSVCSL 361
Qy 356 EDPAHFTSKQSQCLHNQRLDPFFKQOAVCGNAKLEAGBECDCGTQDCCALGETCCDI 415
Db 362 DDFKYISSHNGLTCLQTNPLEMPTYTORRTICGNLLEGGBECDCGNKDNCT--HKLCCDA 419
Qy 416 ATCRFKAGSCAEGPCC-ENCLFMSKERMCRPSPE-ECDLPEYVCGSSASCENHYVOTG 473
Db 420 LTCRLKDNAGCGSDCCCKCKFPKANTICRKSVDVECDTFECFNGSYPPCLLDITYVRDG 479
Qy 474 HPCGLNQMICIDGCMGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQ 533
Db 480 EYCDSGGAFQCGRCRTFDKQCDLIGRSGGAPFCFYDEINTRGDKFGCGTE-----Y 534
Qy 534 CEADNLOGKLIKYVCKFLQIPRATIIYANISGHLCIAVEPASDHA----- 581
Db 535 CLFQHLKGLVCTWEHKDLISRPNLISYIAHVRDQTCVSTYLPSPKPPVASTVSKTSY 594
Qy 582 ----DSQMKWIKDGTSCGSKNVCNRQCVSSSYLYGDCVT---DKNDRGVCNKKKHCH 633
Db 595 YSVDDRDETFFQDGSVCGPDMYCFKMKCKHVFU-MDFETCEASIECSGHGICNNFNHCH 653
Qy 634 CSASYLPDCSVSDLPWGSIDSGNPPVPAIPARLPERRYENIYHSKPMRPPFFLIFP 693
Db 654 CEXGYNPPHCKPKEAF--GSTDDGLVPA-----EKSYMEGRHAPQKQRFQLI-- 702
Qy 694 FFIIFCVLI---ALMVKNVQKRW-----TEDYSSDEQPESESEPK 733
Db 703 FYISLPVLIITAILIK---RKRLCYRGTESESSVSQ-ESSNSK 747

RESULT 4
S47645
tMDC I protein - crab-eating macaque
C;Species: Macaca fascicularis (Crab-eating macaque)
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: S47645
R;Barker, H.L.; Perry, A.C.F.; Jones, R.; Hall, L.
Biochim. Biophys. Acta 1216, 429-431, 1994
A;Title: Sequence and expression of a monkey testicular transcript encoding tMDC I, a nd
A;Reference number: S47645; MUID:94325353; PMID:8049267
A;Accession: S47645
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-736 <BAR>
A;Cross-references: UNIPROT:Q28482; EMBL:X76637; NID:9535016; PID:CAA54085.1; PID:95350
C;Superfamily: mouse meltrin alpha; disintegrin homology
F;392-477/Domain: disintegrin homology <DIS>

Query Match 34.3%; Score 1365.5; DB 2; Length 736;
Best Local Similarity 36.7%; Pred. No. 1.1e-85;
Matches 281; Conservative 145; Mismatches 270; Indels 69; Gaps 19;

Qy 3 VLLSLGLGLRDSNFDSPV-QITVPEKIRSIKEG--IESQASYKIVIEGKPYTVNL 59
Db 4 LLLLSGLRLTAGHSETFVLQITVPRIQIWNVDGVDSETHATYSIKIEGKPYTVLL 63
Qy 60 MQKNFLPHNFRVSYSGTGIMKPLDQDF-QNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQ 118
Db 64 EKQSFLPHFLVLYXNESGLY-VDSFSKGHCFFYQGVADIPKSAVTLATCSGLRGLLQ 122
Qy 119 FENVSYGIEPLESVGFEHVIYQVKH-----KKADVSLYNEK-----DIESRD 161
Db 123 LDNISYGIEPLESPSYEHVVYKNDIAIGHFSFOENYPVAYQIDOSYRILVKRDSNSGA 182
```

```
Qy 162 LSFKLOSAEPQDFAKYIEMHIVIVEKOLYNHMGSDTTTVAQKVQLIGLTNAIFVSNIT 221
Db 183 MLSK-----RTLKIIMDKAMYDMGSAVAVAEKVFQIEGLINTMFSQLNIT 231
Qy 222 IILSSLELWIDENKIATTGEANELHFTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATP 281
Db 232 VMLSSLEIWDNQNKISTGSHADILQRLFLWKQLLQFQRSHDMTYLLIYRNHSTYVGATY 291
Qy 282 QGKMCNDANYAGGVVLPRTISLESIAVLAQLLSLMSGITYDDINKCOCGAVCIMPEA 341
Db 292 HGMACDPKATGATLVPKLTVEAFSVMAQLLGLNGLTYDDIYNICYCGPCTCIMNPD 351
Qy 342 IHFSGVKIFNSCFEDFAHFISKQSQCLHNQRLDPFFKQOAVCGNAKLEAGBECDCGT 401
Db 352 IRSHGMKFFPSSCSMDPEFKQIVLQPELKCFODKTISEMTYQKSAATCGNGILEPTEQDCGY 411
Qy 402 EQCALIGETCCDIATCRFKAGSCAEGPCCEN--CLFMSKERMCRPSFECDCPLPEYV 459
Db 412 RKACTP--KKCCNPADCTLIGFAECGSGPCNNKTCTIYARGHVCRKSIDMCDPFYCNG 469
Qy 460 SSASCENHYVOTGHPCGLNQWICIDGCMGDKQCTDTFGKEVEFGPSECYSHLNSKTD 519
Db 470 TSEFCVPDKAADLEYCSNKTSCYCFKGIHDRDRQCSQLFGKSKSANILCTBEVFNQD 529
Qy 520 VSGNCGISDSGYTQCEADNLQCGKLIKYVCKFLQIPRATIIYANISGHLCIAVEPASD 579
Db 530 KFGNCG-----QRCPPFDILCGKIVCHWTHSELVPMVTDLDIQTYLGGHVCLSAH-ARN 582
Qy 580 HADSQMKWIKDGTSCGSKNVCNRQCVSSSYLYGDCVTDKNDRGVCNKKKHCHCSASYL 639
Db 583 ASKQLGTYTDEDITACGQKQVCRSGRCRYLSELNITKCTTNCGQNGICNDRFHCOCDPYA 642
Qy 640 PPDC--SVQSDLPWGSIDSGNPPV--AIPARLPERRYENIYHSKPMRPPFFLIFPFF 695
Db 643 PPDCDESMSS---PGSIDDGFWLSVDKSVPL-FPKQRAAPK---NNGLLISFYIFSPLL 695
Qy 696 IIFCVLIIMVKNVQKRW-----TEDYSSDEQPESESES 730
Db 696 ILTIAI---VSLKWKMKRFWSKVGTVSSRSRISSED-SSSNNSQSQS 736

RESULT 5
S18968
cyttestin precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: I48784; S18968
R;Senftleben, A.; Wallat, S.; Lemaire, L.; Heinlein, U.A.O.
Dev. Growth Differ. 36, 49-58, 1994
A;Title: Pre and postmeiotic germ cell specific expression of Taz83, a gene encoding a p
A;Reference number: I48784
A;Accession: I48784
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-823 <RES>
A;Cross-references: UNIPROT:Q62287; EMBL:X64227; NID:954364; PID:954265
C;Genetics:
A;Gene: Taz83
C;Superfamily: mouse meltrin alpha; disintegrin homology
F;395-480/Domain: disintegrin homology <DIS>

Query Match 32.4%; Score 1292; DB 2; Length 823;
Best Local Similarity 34.2%; Pred. No. 1.4e-80;
Matches 256; Conservative 156; Mismatches 303; Indels 34; Gaps 15;

Qy 4 LFL-LSGLGG-LRMDSNFDSPVQITVPEKIRSIKEG--IESQASYKIVIEGKPYTVNL 59
Db 5 LFLVLSVGLGVIAAGKDVTFLLQITVPEKIDTNIQDAKEATQVTVVTVIEGKAYTLQL 64
Qy 60 MQKNFLPHNFRVSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQ 119
Db 65 EKQSFLPHFLFGTYLRDKLGLTQYFSLVKTHCFYQGHAAEIPVSTVLTSTCSGLRGLLQ 124
```





```
QY 368 QCLHNPRLDPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGSNCA 427
D 369 FAALVHMLHGLGHOHDH-SACFCEKHFCLMHNITKESG---FSCSSDYFHOFLE 424
QY 365 QKSCQCLHNPRLDPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGS 424
D 425 HKGACLFNPRGRKRDRSACNGVVEDEECDCGSA--CHL--DPCCD-PTCTLKEHA 479
QY 425 NCAEGPCCNCLFMSKERMCRPSFECEDLPYCNCGSSASCPENHYVQTGHPGCLNQCWICI 484
D 480 ECHGLCLDCTFRKGFCLCRPTQDECDLPYCGSSAECPADSKQDGLCDRIHY-CS 538
QY 485 DGVCMGDKQCTDTFGKEVEFGSECYSHLNSKTDVSGNCG---ISDSGYTQCEADNLCQ 541
D 539 GGCKNPDNQCVMNIYGPARSAPEDCYISMNTRGDRFGNCGHPTEDQOYVTVCSDDNVFC 598
QY 542 GKLIKVVGKFLQIPRA---TIIYANISGHLCIAVEF--ASDHADSOQRMWKDGTSCG 595
D 599 GKLICITGV---QSLPRVKAQHTVIOVPHDNDWCMSMDADNITDTPDNGNVHV--GTSCA 652
QY 596 SNKVCNRQCVSSSYLYGDTCTDK-CNDRGVGNKKHCHCSASYLPPDCVQSDLMWPGGS 654
D 653 PNKVCCTDYSVHHSILLYDCRPEESCHGKVCNLRHCHCESGFAPDCKNPGN---GGS 709
QY 655 IDSGNFPFVAIPARLPERRYENIYHSPMRP-----PFLFPIFFIIFCVLIAIMVKVN 709
D 710 VDSG---PFGMQVTNNSGESIARGQSLRQDVYKLVLLVPLFLV--LLLCSLLTIS 764
QY 710 FORKKWRT-----EDYSSDEQPESE 729
D 765 YLCSEVQTAVAEVESSTETLESE 789

RESULT 10
S55059
fertilin alpha-I - crab-eating macaque
C/Species: Macaca fascicularis (Crab-eating macaque)
C/Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C/Accession: S55059
R/Perry, A.C.F.; Gichuhi, P.M.; Jones, R.; Hall, L.
Biochem. J. 307, 843-850, 1995
A/Title: Cloning and analysis of monkey fertilin reveals novel alpha subunit isoforms.
A/Reference number: S55059; MUID:95260313; PMID:7741716
A/Accession: S55059
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-905 <PR>
A/Cross-references: UNIPROT:Q28476; EMBL:X79808; NID:G794072; PIDN:CAAS6203.1; PID:G7940
C/Superfamily: mouse meltrin alpha; disintegrin homology
F/442-522/Domain: disintegrin homology <DIS>
F/376/Active site: Glu #status predicted

Query Match 25.7% Score 1022.5; DB 2; Length 905;
Best Local Similarity 33.6%; Pred. No. 4.4e-62;
Matches 250; Conservative 123; Mismatches 297; Indels 75; Gaps 29;

QY 22 LPQVITYPEKIRSIKEGISOASKYKIVIEKPYTNL-WQKNFLPHNFRVYSYSGTGM 80
D 83 IPKRLTVGGDSSV--EGL-----SYLLFMQKQHLVKKRSHFVNFPVSYH-NGIL 135
QY 81 KPLDQDQFQCHYQGYEYKPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVY 140
D 136 QGSPFISHDCHYEYIEGYGSGFVSVNTCAGLIGLILKEEKSYSIEPMOSSRRFHEVLY 195
QY 141 QVGHKKADVSLYNEKDIETSRD-----LSFKLQSAEPQO-----DPKAYIEMHIVIEK 187
D 196 TMAH-EARVSC---GVTSRDSHVSTVSWQGRSKPHDLQALSILWSHTKYEMFVVVNN 250
QY 188 QLYNHMGSDITVAQKVPOLIGLITNALFVSNITILSSLELWIDENKIATTEGANELH 247
D 251 QRFQMGSNINETQVRVVDVIALANGFTRGINTVWLAGMEIWEGLDIDVAVDLQITLR 310
QY 248 TFLRWKTSYLVR-PHVDVALLVVRKSNVYCATFGQKMCDCDANYAGGV-VLHPRTISLES 305
D 311 NFNEWRQEMLFRRAKHDVAHVHGHFGQNTGQAFLSGACSSGFAAAVESFHEDMLL-- 368
```

```
QY 306 LAVILAQLLSLNGITYDDINKQC--SGAVCMNPNPAIHFSGVKVIKFSNCSFEDFAHFISK 364
D 369 FAALVHMLHGLGHOHDH-SACFCEKHFCLMHNITKESG---FSCSSDYFHOFLE 424
QY 365 QKSCQCLHNPRLDPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGS 424
D 425 HKGACLFNPRGRKRDRSACNGVVEDEECDCGSA--CHL--DPCCD-PTCTLKEHA 479
QY 425 NCAEGPCCNCLFMSKERMCRPSFECEDLPYCNCGSSASCPENHYVQTGHPGCLNQCWICI 484
D 480 ECHGLCLDCTFRKGFCLCRPTQDECDLPYCGSSAECPADSKQDGLCDRIHY-CS 538
QY 485 DGVCMGDKQCTDTFGKEVEFGSECYSHLNSKTDVSGNCG---ISDSGYTQCEADNLCQ 541
D 539 GGCKNPDNQCVMNIYGPARSAPEDCYISMNTRGDRFGNCGHPTEDQOYVTVCSDDNVFC 598
QY 542 GKLIKVVGKFLQIPRA---TIIYANISGHLCIAVEF--ASDHADSOQRMWKDGTSCG 595
D 599 GKLICITGV---QSLPRVKAQHTVIOVPHDNDWCMSMDADNITDTPDNGNVHV--GTSCA 652
QY 596 SNKVCNRQCVSSSYLYGDTCTDK-CNDRGVGNKKHCHCSASYLPPDCVQSDLMWPGGS 654
D 653 PNKVCCTDYSVHHSILLYDCRPEESCHGKVCNLRHCHCESGFAPDCKNPGN---GGS 709
QY 655 IDSGNFPFVAIPARLPERRYENIYHSPMRP-----PFLFPIFFIIFCVLIAIMVKVN 709
D 710 VDSG---PFGMQVTNNSGESIARGQSLRQDVYKLVLLVPLFLV--LLLCSLLTIS 764
QY 710 FORKKWRT-----EDYSSDEQPESE 729
D 765 YLCSEVQTAVAEVESSTETLESE 789

RESULT 11
S71949
metalloproteinase 12 (EC 3.4.24.-) precursor - human
A/Alternate names: disintegrin 12; myeloma cell metalloproteinase MCMP
C/Species: Homo sapiens (man)
C/Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 09-Jul-2004
C/Accession: S71949; PC4264
R/McKie, N.; Dallas, D.J.; Edwards, T.; Apperley, J.F.; Russell, R.G.G.; Croucher, P.I.
Biochem. J. 318, 459-462, 1996
A/Title: Cloning of a novel membrane-linked metalloproteinase from human myeloma cells.
A/Reference number: S71949; MUID:96404892; PMID:8809033
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-660 <MCK>
A/Cross-references: UNIPROT:Q13443
A/Experimental source: myeloma cells
R/McKie, N.; Edwards, T.; Dallas, D.J.; Houghton, A.; Stringer, B.; Graham, R.; Russell,
Biochem. Biophys. Res. Commun. 230, 335-339, 1997
A/Title: Expression of members of a novel membrane linked metalloproteinase family (ADAM)
A/Reference number: PC4263; MUID:97168971; PMID:9016778
A/Accession: PC4264
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-337 <MCF>
A/Experimental source: articular chondrocyte
C/Function:
A/Description: involved in cell/cell and cell/matrix interaction in a variety of cell ty
A/Note: membrane-bound; belongs to reprotolysin family of metalloproteinases
C/Superfamily: mouse meltrin alpha; disintegrin homology
C/Keywords: hydrolase; metalloproteinase; zinc
F/1-24/Domain: signal sequence #status predicted <SIG>
F/25-660/Product: pre-metalloproteinase 12 #status predicted <MAT>
F/54-59/Region: autoinhibitory
F/295-378/Domain: disintegrin homology <DIS>
F/622-642/Domain: transmembrane #status predicted <TM1>
F/57,229,233,239/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status i
F/229,233,239/Binding site: zinc, catalytic (His) (active) #status predicted
F/230/Active site: Glu #status predicted
```

```

Query Match      25.5%; Score 1016.5; DB 2; Length 660;
Best Local Similarity 34.7%; Pred. No. 7.6e-62;
Matches 216; Conservative 109; Mismatches 207; Indels 91; Gaps 16;

Qy      105 VMVSTCTGLRGVLPENVSVCIEPLESSVGFHVIYQV-----KHKADVSLNEKDIES 159
      : : : : | | | | | | | | | | | | | | | | : | | | |
Db      10 LLLATVLDLRLGLHLENASYIEPLQSSHFHEHIYRMDVDYKEPLKCGVS---NKDIEK 66
      : : : : | | | | | | | | | | | | | | | | : | | | |

Qy      160 R-----DLSFKLQSAEPQODFAKYIEMHVIVEKOLYNHNGSDTTVVAAQVKFQ 206
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      67 ETAKDEEEPPSMTQLRRRAVLQP---TRYVELFIVVDKERYDDMGRNQTAVEREMIL 123
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |

Qy      207 LIGLITNAIFVSNITIISSLELWIDENKIATGEEANELLHTFLRWKTSYLVLI-RPHDVA 265
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      124 LANYLDSWYIMLNIIRIVLGLIEWTNGNLINIVGGAGDVLGNFQWREKELITRRHDSA 183
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |

Qy      266 FLLVYRKSNNVGTATPGKMCNDANYAGGVVLHPRTISLES LAVILAOLLSLWSGITVDDI 325
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      184 QLVLLKGGFGGTAGMAFVGTVCSSHAGGINVFGQ-ITVETFASIVAHELGHNLMGNHDDG 242
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |

Qy      326 NKOCSCGAVCIMPEAHFHSVKLTFSCNSPEDFAHFTSKQKSOCLHNQPRLDPPFKQOAV 385
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      243 RDCSCGAKSCIMNSGA---SGSRNFSSCSAEDFEKLTNLNKGNCLLNIPKPDSEAYSAPS- 298
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |

Qy      386 CGNAKLEAGEECDCGTBQDCALITGETCCDIATCRFKAGSNCAEGPCCECCLFWSKERMCR 445
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      299 CGNKLVDAGEECDCGTPEKCEL--DPCCEGSTCKLKSFAECAYGDCCKCRFLPGGTLCR 356
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |

Qy      446 PSPEECDLPRYCNCGSSASCENPHYVGTGHPCLGNQWITCIDGVCMSGDKQCTDFGKEVEF 505
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      357 GKTSECDVPEYCNCGSSQFCQDPVFIQNGYFCQNNKAYCYNGMCQYYDAQCVIFGSKAKA 416
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |

Qy      506 GPSECYSHLNSKTDVSGNCIGSIDSGYTCQE----- 535
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      417 APKDCFTIEVNSKGRDFNGCFSGNEYKKCATGLSLKFHAPFLSTMLQEAVRQTGTVLGGS 476
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |

Qy      536 -----ADNLQCKGLKCKYVGKFLLOIPRATIIYANIS-----GHLCIAVEF-- 576
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      477 VCKWKSDCRIVTLVKNALCKLQCENV-----QEIPVEGIVPAIIQTSPRTCKCWGVDFQL 532
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |

Qy      577 ASDHADSQKMWIKDGTSCGSKNVCRNORCVSSSYLVGYDCTTD-KCNDRGVCNKKKCHCS 635
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      533 GSDVPDPG--WVNEGTCGAGKICRNFCQYDASVLANYDCDVQKCKCHGVGCNKNKCHCE 590
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |

Qy      636 ASYLPPDCSVQSDLWPGGSDSG 658
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      591 NGWAPPNCETKG---YGGSVDSG 610
      : : : : | : : : | : : : | : : : | : : : | : : : | : : : |

```

```

RESULT 12
S60257
meltrin alpha - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S60257
R:Yagami-Hironaka, T.; Sato, T.; Kurisaki, T.; Kamijo, K.; Nabeshima, Y.I.; Fujisawa-Sehara, Y.
Nature 377, 652-656, 1995
A:Title: A metalloprotease-disintegrin participating in myoblast fusion.
A:Reference number: S60257; MUID:9026308; PMID:7566181
A:Accession: S60257
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-903 <YAG>
A:Cross-references: UNIPROT:Q61824; EMBL:D50411; NID:G1054586; PIDN:BAA08912.1; PID:G1054586
C:Superfamily: mouse meltrin alpha; disintegrin homology
F:421-503/Domain: disintegrin homology <DIS>
F:349/Active site: Glu #status predicted

```

Query Match 21.5%; Score 857; DB 2; Length 903;  
Best Local Similarity 30.7%; Pred. NO. 9.4e-51;  
Matches 199; Conservative 115; Mismatches 275; Indels 60; Gaps 19;

Qy	47	KIVIEGKPYTNL-MQKNFLPHNFRVYSGTIGIMKPLDQDDQNFQCHGYGIEGPKSV	105
Db	72	QQLSRLILLSERNEGLIANGFTETHYLQDQDVTSLRNRHDCYHGHVQGDAASV	131
Qy	106	MVSTCTGLRGVLQFENVSGTEPLESSYGFEBHVIQVXKHKADVSL-----YNEKDE	158
Db	132	SLSTCSDLRGLIMENKTYISLEPMKNTD---SYKLVPASMTNIOGLCSQHKNLSLT	187
Qy	159	SRDLS---FKLOSABPOOD---FAKYIEBHVIVEKQLYNNHMGSDTTVVAQKVQFQILGLTN	212
Db	188	MEDVSPGTSQMARRRHKRETLKMTKYVELVIVADNREFQOGKOLEKVQRKRLIEIANHVD	247
Qy	213	AIFVSNFTIITILLSLEWIDENKATTCGEANELLHTELRM-KTSLVLVRPHDVAELLVYR	271
Db	248	KFYRPLNIRIVLGVVEVWMDIDKCSISODPTRLHBFELDMRKIKULPRKSHDNAQLI---	304
Qy	272	EKSNYVGAFTQG-----KMCDAANYAGGVVLHPRTISLESIAVILAQLLSLSMGITY	322
Db	305	-----SGVYFGTTIGMAPINSMCTAEQSGGVWMDHSDSPL-GAAVTLAHELGHNFGMNH	358
Qy	323	DDINK-COCGAV----CINMPEAITHFGSVKIFNSCFEDFAHFTSKOKSQCILHNQPRLD	377
Db	359	DTLERGSCRMABKGGCINNP-STGFPPFPMVFSGSRKOLEASLEKMGMCMLFNLPEVK	417
Qy	378	PFEKQAAVCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLF	437
Db	418	QAFGGRK-CGNGYVEEGEBCDCGPEECT--NRCNATTCTLKPDAVCAHGQCCEDCOL	473
Qy	438	MSKERMCRPSPEBCDLPEYCNNGSSASCSPENHYVQTGHPHPCGLNQWITCIDGVCMGSDKQCTD	497
Db	474	KPGTACRGSSNSCDLPEFCGTGTAPHCPANVYLHDGHPFCQGVGVCYNGICQTHEQQCVT	533
Qy	498	TFGKEVEFGPSECYSHLNSKTDVSGNCG-ISDSGYTOQCEANLQCGKILCKYVGFGLQI	556
Db	534	LWPGKAPAPGICFERVNSAGDPYNGCKGSKSAFAKCELRDAKCGKICQCGGASRPVIG	593
Qy	557	PRATIIYANISGH-----LC--IAVEPASDHDASQJMWIKDGTSCGSKNVCNRORCVSSS	609
Db	594	TNANSIETNIPQBSGRILCRGTHVYLGDDNPDPG--LVLAGTKCAEGKICLNRRQCNIS	651
Qy	610	YLGYDCTTDKCDRGVCKNKKHCHCSASYLPDPCSVQSDLWPGGSDISG	658
Db	652	VFGVHKCAMOCHGRGVCKNRRKNCHCEAHWAPFC---DKFGFGGSDISG	697

RESULT 13

I49281

fertilin alpha precursor - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:species: mus musculus (mouse house)  
C:Date: 02-Jul-1996 #sequence revision 02-Jul-1996 #text change 16-Aug-2004

C;Accession: I49281

R:Wolfsberg, T.G.: Straight. P.D.: Gerena. R.L.: Huovila. A.P.: Primakoff. P.: Myles. D.

Dev. Biol. 169, 378-383, 1995

A;Title: ADAM, a widely

A:Reference number: I48100; MUID:95269891; PMID:7750654

A;Accession: I49281

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-600 <RES>

A;Cross-references: EMBL:U22056; NID:g965009; PIDN:AAA74920.1; PID:g965010

**C;Genetics:**

A;Gene: ADAM 1

**C; Superfamily: disintegrin homology**

F;246-326/Domain: disintegrin homology <DIS>

F;180/Active site: Glu #status predicted

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

Query Match 20.0%; Score 795; DB

Best Local Similarity 31.5%; Pred. No. 1e-46;  
Matches 102; Concentrations 92; Mismatches 256; Indels

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

QY 161 DLSFKLSAEPQQ-----DFAKYIEMHVI VEKQLYNHMGSDTTVVAQKVFQLIGLTN 21







Db	197	FRFVELVVDKAWVTNRNGDLDKIKTRMYEIVNTVNEIYRYMYIHVALVGLIWSNEDK	256
Qy	236	IATTGEANEILLHTFLRW-KTSYLVLRPHDVAFLVYREKSNYYVGATFQGRMCDANYAGGV	294
Db	257	ITVKPEAGYTLNAFGWRKTDLLTRKHGDNAAQLLTAIDLDRVIGLAVVGSMSCHPKRSTGI	316
Qy	295	VLHPRITISLESILAVILAQLLSLMSGITYDINKQCSGAVCIIMNPEAIHFSGVKIFSNCS	354
Db	317	IODYSEINL-VVAVIMAHMGHNLGINHDS-GYCSGDIYACIMRPE-ISPEPSTFFSNCS	373
Qy	355	PEDPAHETISKQSOCLHNPRLDPFFKQAVCGNAKLEAGEECDGTEQOCALIGETCCD	414
Db	374	YFECWDFIMNHNPECILNEP-LGTDIISPVCNGLLEVGEECDGCPENCQ---NECCD	429
Qy	415	IATCRFKAGSNCAEGPCCNCLFMSKERMCRPSFEECDLPEYCNCGSSASCENHYVQTGH	474
Db	430	AATCKLKSQSGCHGDCCEQCKFSKSGTECRASMECDPAEHCTGQSSECPADVPHKNGQ	489
Qy	475	PCGLNQWICIDGVCMGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQC	534
Db	490	PCLDNYGYCNGNCPIMYHQCYDLFGADVYEAEADSCPER-NQKGNYYGYCRKENGKIPC	548
Qy	535	EADNLQCGKLIQKVVYVGFLLQIPRATIIYANISGHLCIAVEFASDHADSKMWIKGTSC	594
Db	549	APEDVKCGRLYCK-----DNSPGQNNPCRMFYSNE-DEHKGWVLPQTKC	591
Qy	595	GSNKVCRNQRVCV	606
Db	592	ADGKVCNSGHCV	603

Search completed: January 10, 2005, 22:28:33  
Job time : 45 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:04:22 ; Search time 201 Seconds  
(without alignments)  
2101.119 Million cell updates/sec

Title: US-10-054-683-19

Perfect score: 3984

Sequence: 1 MWVFLSLGLGURMDSNF.....WRTEDYSSDRQPSSEPKG 734

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3369.5	99.6	735	1 AD02_HUMAN	Q9965 homo sapien
2	3690.5	92.6	735	1 AD02_MACFA	Q28478 macaca fasc
3	3650.5	76.9	579	2 Q6P2G0	Q6P2G0 homo sapien
4	3065.5	76.9	579	2 AAH64547	Aah64547 homo sapi
5	2820	70.8	735	2 Q866A8	Q866A8 sus scrofa
6	2723	68.3	745	1 AD02_BOVIN	Q77780 bos taurus
7	2685	67.4	751	1 AD02_RABIT	Q28660 oryctolagus
8	2479.5	62.2	737	1 AD02_RAT	Q63202 rattus norv
9	2453.5	61.6	735	1 AD02_MOUSE	Q60718 mus musculu
10	2326	58.4	735	1 AD02_CAVPO	Q60411 cavia porce
11	1586.5	39.8	746	1 AD18_MACFA	Q95134 macaca fasc
12	1582.5	39.7	739	1 AD18_HUMAN	Q9Y3G7 homo sapien
13	1524.5	38.3	787	2 Q8TC27	Q8TC27 homo sapien
14	1524.5	38.3	787	2 AAQ89099	Aaq89099 homo sapi
15	1519.5	38.1	787	2 Q8TC42	Q8TC42 homo sapien
16	1493.5	37.5	715	2 Q6UXJ9	Q6UXJ9 homo sapien
17	1493.5	37.5	715	2 AAQ88687	Aaq88687 homo sapi
18	1437.5	36.1	750	2 Q6P901	Q6P901 mus musculu
19	1437.5	36.1	750	2 AAH60983	Aah60983 mus muscu
20	1433	36.0	750	2 Q8K410	Q8K410 mus musculu
21	1383.5	34.7	756	2 Q28483	Q28483 macaca fasc
22	1371.5	34.4	730	2 Q80Y08	Q80Y08 mus musculu
23	1367.5	34.3	677	2 Q8BJ80	Q8BJ80 mus musculu
24	1366.5	34.3	719	1 AD18_MOUSE	Q9R157 mus musculu
25	1365.5	34.3	736	2 Q28482	Q28482 macaca fasc
26	1293	32.5	822	2 Q62287	Q62287 mus musculu
27	1292	32.4	740	2 P70534	P70534 rattus norv
28	1288	32.3	777	2 Q60472	Q60472 cavia porce
29	1237	31.0	771	2 Q8CDV5	Q8CDV5 mus musculu
30	1237	31.0	771	2 Q9DAF0	Q9DAF0 mus musculu
31	1211	30.4	819	1 AD09_HUMAN	Q13443 homo sapien

32	1206	30.3	751	2	Q60816	Q60816 mus musculu
33	1166	29.3	845	1	AD09_MOUSE	Q61072 mus musculu
34	1161.5	29.2	873	2	O42595	O42595 xenopus lae
35	1059	26.6	919	2	Q28659	Q28659 oryctolagus
36	1042	26.2	792	2	O19061	O19061 saquinus oe
37	1041	26.1	812	2	O77779	O77779 bos taurus
38	1031.5	25.9	825	2	Q28477	Q28477 macaca fasc
39	1028.5	25.8	791	1	AD1A_MOUSE	Q60813 mus musculu
40	1024.5	25.7	789	1	AD01_RAT	P70505 rattus norv
41	1022.5	25.7	905	2	Q28476	Q28476 macaca fasc
42	1016	25.5	821	2	O19060	O19060 saquinus oe
43	1015.5	25.5	825	2	O46652	O46652 papio anubi
44	1014.5	25.5	838	2	O19056	O19056 papio anubi
45	1002.5	25.2	804	2	Q60410	Q60410 cavia porce

## ALIGNMENTS

RESULT 1

ID	AD02_HUMAN	STANDARD;	PRT;	735 AA.
AC	Q9965; P78326; Q9UQ08;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)			
DE	(Fertilin beta subunit) (PH-30) (PH30).			
GN	Name=ADAM2; Synonyms=FTNB;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Testis;			
RX	MEDLINE=97193554; PubMed=9041139;			
RA	Vidaeus C.M., von Kap-Herr C., Golden W.L., Eddy R.L., Shows T.B.,			
RA	Herr J.C.;			
RT	"Human fertilin beta: identification, characterization, and			
RT	chromosomal mapping of an ADAM gene family member.";			
RL	Mol. Reprod. Dev. 46:363-369 (1997).			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Testis;			
RX	MEDLINE=96295488; PubMed=8702389;			
RA	Gupta S.K., Alves K., Palladino L.O., Mark G.E., Hollis G.F.;			
RT	"Molecular cloning of the human fertilin beta subunit.";			
RL	Biochem. Biophys. Res. Commun. 224:318-326 (1996).			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Testis;			
RX	MEDLINE=97224507; PubMed=9070941;			
RA	Burkin H.R., Burkin D.J., Davey P.M., Griffin D.K., Affara N.A.;			
RT	"Mapping, sequence, and expression analysis of the human fertilin beta			
RT	gene (FTNB).";			
RL	Genomics 40:190-192 (1997).			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Testis;			
RA	Hall L., Frayne J.;			
RT	"Nucleotide sequence of the human fertilin beta transcript.";			
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Testis;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T.I., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			

RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Basak S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallic D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -1- FUNCTION: Sperm surface membrane protein that may be involved in  
 CC sperm-egg plasma membrane adhesion and fusion during  
 CC fertilization. Could have a direct role in sperm-zona binding or  
 CC migration of sperm from the uterus into the oviduct. Interactions  
 CC with egg membrane could be mediated via binding between its  
 CC disintegrin-like domain to one or more integrins receptors on the  
 CC egg. This is a non catalytic metalloprotease-like protein.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC -1- Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q99965-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q99965-2; Sequence=VSP\_005471;  
 CC Note=No experimental confirmation available;  
 CC -1- TISSUE SPECIFICITY: Expressed specifically in spermatogenic cells  
 CC in the seminiferous cells. Not detected in fetal tissues.  
 CC -1- DOMAIN: A tripeptide motif (PEE) within disintegrin-like domain  
 CC could be involved in the binding to egg integrin receptor and thus  
 CC could mediate sperm/egg binding.  
 CC -1- FM: The prodomain and the metalloprotease domain are cleaved  
 CC during the epididymal maturation of the spermatozoa.  
 CC -1- MISCELLANEOUS: In mammals, exists as a heterodimer composed of an  
 CC alpha and beta subunits. In human, fertilin alpha is a pseudogene.  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin domain.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U52370; AAC51110.1; -;  
 CC EMBL: U38805; AAD04206.1; -;  
 CC EMBL: X93974; CAA67753.1; -;  
 CC EMBL: AJ133005; CAB40813.1; -;  
 CC EMBL: BC034957; AAH34957.1; -;  
 CC PIR: JC4861; JC4861.  
 CC HSP: P30403; IN4Y.  
 CC MEROPS: M12.950; -;  
 CC Genew: HGNC:198; ADAM2.  
 CC MIM: 601533; -;  
 CC GO: GO:0005887; C: integral to plasma membrane; TAS.  
 CC GO: GO:0005178; F: integrin binding; TAS.  
 CC GO: GO:0008237; F: metalloproteinase activity; TAS.  
 CC GO: GO:0007342; P: fusion of sperm to egg plasma membrane; TAS.  
 CC InterPro: IPR006586; ADAM\_cysteine.  
 CC InterPro: IPR001762; Disintegrin.  
 CC InterPro: IPR006209; EGF like.  
 CC InterPro: IPR001590; Peptidase M12B.  
 CC InterPro: IPR002870; Peptidase M12B\_N.  
 CC InterPro: IPR006025; Pept\_M\_Zn\_BS.  
 CC Pfam: PF00200; Disintegrin; 1.  
 CC Pfam: PF01562; Pep M12B propep; 1.  
 CC Pfam: PF01421; Reptolysin; 1.

DR PRINTS; PR00289; DISINTEGRIN.  
 DR ProDom; PD000664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR PROSITE; PS00215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; 1.  
 DR PROSITE; PS00214; DISINTEGRIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; FALSE NEG.  
 DR PROSITE; PS01186; EGF\_2; FALSE NEG.  
 KW Alternative splicing; Cell adhesion; EGF-like domain; Glycoprotein;  
 KW Signal; Transmembrane.  
 FT SIGNAL 1 16 Potential.  
 FT PROPEP 17 174  
 FT CHAIN 175 735 ADAM 2.  
 FT DOMAIN 175 686 Extracellular (Potential).  
 FT TRANSMEM 687 707 Potential.  
 FT DOMAIN 708 735 Cytoplasmic (Potential).  
 FT DOMAIN 175 382 Metalloprotease-like.  
 FT DOMAIN 383 475 Disintegrin-like.  
 FT DOMAIN 477 606 Cys-rich.  
 FT DOMAIN 612 645 EGF-like.  
 FT DISULFID 287 370 By similarity.  
 FT DISULFID 329 354 By similarity.  
 FT DISULFID 445 458 Potential.  
 FT DISULFID 616 627 By similarity.  
 FT DISULFID 621 633 By similarity.  
 FT DISULFID 635 644 By similarity.  
 FT CARBOHYD 122 122 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 220 220 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 353 353 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 459 459 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 566 566 N-linked (GlcNAc...) (Potential).  
 FT VARSPLIC 172 190 Missing (in isoform 2).  
 FT FTid=VSP\_005471.  
 FT CONFLICT 3 3 Missing (in Ref. 2).  
 FT CONFLICT 21 21 D -> H (in Ref. 3).  
 FT CONFLICT 99 99 E -> D (in Ref. 3).  
 FT CONFLICT 106 106 V -> G (in Ref. 3).  
 FT CONFLICT 170 170 V -> A (in Ref. 2).  
 FT CONFLICT 288 288 D -> H (in Ref. 1).  
 FT CONFLICT 321 321 I -> T (in Ref. 1).  
 FT CONFLICT 388 388 G -> S (in Ref. 3).  
 FT CONFLICT 396 398 EEC -> DEF (in Ref. 3).  
 FT CONFLICT 501 501 G -> S (in Ref. 3).  
 FT CONFLICT 529 529 D -> Y (in Ref. 3).  
 FT CONFLICT 579 579 S -> G (in Ref. 3).  
 FT CONFLICT 588 588 M -> L (in Ref. 3).  
 FT CONFLICT 603 603 N -> D (in Ref. 3).  
 FT CONFLICT 629 630 NK -> KQ (in Ref. 3).  
 FT CONFLICT 638 638 S -> F (in Ref. 3).  
 SQ SEQUENCE 735 AA; 82456 MW; 92867B5340BEE01F CRC64;

Query Match 99.6%; Score 3969.5; DB 1; Length 735;  
 Best Local Similarity 99.7%; Pred. No. 1.2e-261;  
 Matches 733; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MW-VLFLSLGLGRMDSNFDLPVQITVPEKIRSIKEGIESQASQYKIVIEGKPYTVNL 59  
 DB 1 MWVFLSLGLGRMDSNFDLPVQITVPEKIRSIKEGIESQASQYKIVIEGKPYTVNL 60  
 QY 60 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGLQF 119  
 DB 61 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGLQF 120  
 QY 120 ENVSYGIEPLESSVGFHEHYQVKKKADVSLNKKDIESRDLSEFKLQSPQDFAKYI 179  
 DB 121 ENVSYGIEPLESSVGFHEHYQVKKKADVSLNKKDIESRDLSEFKLQSPQDFAKYI 180  
 QY 180 EMHIVIVEKQLYNHMGSDTTVAQKFLIGLTNAIFVSNITILSLSLWIDENKIATT 239  
 DB 181 EMHIVIVEKQLYNHMGSDTTVAQKFLIGLTNAIFVSNITILSLSLWIDENKIATT 240  
 QY 240 GEANELHTFLRWKTSYLVLRPHDVAFLVYREKSNVYGVATFOGKMCNDANAGVVLHPR 299

Db 241 GEANELHTEFLRMKTSYLVLRPHDVAFLVYREKSNVYVATFQGMCDANYAGVWLHPR 300  
 Qy 300 TISLESIAVLAQLLSLSMGTDDINKCQSGAVCIWNPENAIHFGSKVIFNSGSPEDFA 359  
 Db 301 TISLESIAVLAQLLSLSMGTDDINKCQSGAVCIWNPENAIHFGSKVIFNSGSPEDFA 360  
 Qy 360 HFTSKQKQCLHNPRLDPFFKQAVCGNAKLEAGEBCDGTEDCALIGETCCDIATCR 419  
 Db 361 HFTSKQKQCLHNPRLDPFFKQAVCGNAKLEAGEBCDGTEDCALIGETCCDIATCR 420  
 Qy 420 FKAGSNCAEGPCPCNCLFMSKRCRPSRECDLPYCNCGSSASCENHYVQTGHPGCLN 479  
 Db 421 FKAGSNCAEGPCPCNCLFMSKRCRPSRECDLPYCNCGSSASCENHYVQTGHPGCLN 480  
 Qy 480 QWICIDGVCMSGDKQCTDTFGKEVFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNL 539  
 Db 481 QWICIDGVCMSGDKQCTDTFGKEVFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNL 540  
 Qy 540 QCGKLIKYVGVKFLQIPRATIIYANISGHLCIAVEFASDHADSKQWIKDGTSCGNKV 599  
 Db 541 QCGKLIKYVGVKFLQIPRATIIYANISGHLCIAVEFASDHADSKQWIKDGTSCGNKV 600  
 Qy 600 CRNORCVSSVLYGDCVTDKNDRGVGNKKGCHCSASYLPPDCSVQSDLWPGGSDSGN 659  
 Db 601 CRNORCVSSVLYGDCVTDKNDRGVGNKKGCHCSASYLPPDCSVQSDLWPGGSDSGN 660  
 Qy 660 FPPVAIPARIPERYIENIYHSHKPMRPFLLFPFFIIFCVLIAMVKNVFORKKWRTED 719  
 Db 661 FPPVAIPARIPERYIENIYHSHKPMRPFLLFPFFIIFCVLIAMVKNVFORKKWRTED 720  
 Qy 720 YSSDEQPESESEPKG 734  
 Db 721 YSSDEQPESESEPKG 735

## RESULT 2

AD02 MACFA STANDARD; PRT; 735 AA.  
 ID AD02 MACFA  
 AC Q28478; Q28472;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)  
 DE (Fertilin beta subunit) (PH-30) (PH30) (PH30-beta).  
 GN Name=ADAM2; Synonyms=FTNB;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 OX NCBI\_TaxID=9541;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Testis;  
 RC MEDLINE=95260313; PubMed=7741716;  
 RA Perry A.C.F., Gichuhi P.M., Jones R., Hall L.;  
 RT "Cloning and analysis of monkey fertilin reveals novel alpha subunit  
 isoforms";  
 RL Biochem. J. 307:843-850(1995).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=96341733; PubMed=8720115;  
 RA Ramarao C.S., Myles D.G., White J.M., Primakoff P.;  
 RT "Initial evaluation of fertilin as an immunoreceptive antigen and  
 molecular cloning of the cynomolgus monkey fertilin beta subunit.";  
 RL Mol. Reprod. Dev. 43:70-75(1996).  
 CC -!- FUNCTION: Sperm surface membrane protein that may be involved in  
 sperm-egg plasma membrane adhesion and fusion during  
 fertilization. Could have a direct role in sperm-zona binding or  
 migration of sperm from the uterus into the oviduct. Interactions  
 with egg membrane could be mediated via binding between its  
 disintegrin-like domain to one or more integrins receptors on the  
 egg. This is a non catalytic metalloprotease-like protein (By

CC similarity).  
 CC -!- SUBUNIT: Heterodimer with ADAM1/fertilin alpha.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Expressed specifically in testis.  
 CC -!- DOMAIN: A tripeptide motif (FDE) within disintegrin-like domain  
 CC could be involved in the binding to egg integrin receptor and thus  
 CC could mediate sperm/egg binding (By similarity).  
 CC -!- FTM: The signal and the metalloprotease domain are cleaved during  
 CC the epididymal maturation of the spermatozoa (By similarity).  
 CC -!- SIMILARITY: Belongs to peptidase family M12B.  
 CC -!- SIMILARITY: Contains 1 disintegrin domain.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X77653; CAA54733.1; -.  
 CC EMBL; U3959; AAB52699.1; -.  
 CC PIR; G02937; G02937.  
 CC HSP; P30403; IN4Y.  
 CC MEROPS; M12.950; -.  
 CC InterPro; IPR006586; ADAM cysteine.  
 CC InterPro; IPR001762; Disintegrin.  
 CC InterPro; IPR000742; EGF 2.  
 CC InterPro; IPR006209; EGF-like.  
 CC InterPro; IPR001590; Peptidase M12B.  
 CC InterPro; IPR002870; Peptidase M12B\_N.  
 CC Pfam; PF01562; Pep M12B proper; 1.  
 CC Pfam; PF01421; Reptolysin; 1.  
 CC PRINTS; PR00289; DISINTEGRIN.  
 CC ProDom; PD000664; disintegrin; 1.  
 CC SMART; SM00608; ACR; 1.  
 CC SMART; SM00050; DISIN; 1.  
 CC PROSITE; PS00215; ADAM MEPRO; 1.  
 CC PROSITE; PS00427; DISINTEGRIN\_1; 1.  
 CC PROSITE; PS00214; DISINTEGRIN\_2; 1.  
 CC PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 CC PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 CC PROSITE; PS00026; EGF\_3; 1.  
 CC Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
 FT SIGNAL 1 16 Potential.  
 FT PROPEP 17 174 By similarity.  
 FT CHAIN 175 735 ADAM 2.  
 FT DOMAIN 17 686 Extracellular (Potential).  
 FT TRANSMEM 687 707 Potential.  
 FT DOMAIN 708 735 Cytoplasmic (Potential).  
 FT DOMAIN 175 382 Metalloprotease-like.  
 FT DOMAIN 383 475 Disintegrin-like.  
 FT DOMAIN 476 606 Cys-rich.  
 FT DOMAIN 612 645 EGF-like.  
 FT DISULFID 287 370 By similarity.  
 FT DISULFID 329 354 By similarity.  
 FT DISULFID 445 458 Potential.  
 FT DISULFID 616 627 By similarity.  
 FT DISULFID 621 633 By similarity.  
 FT DISULFID 635 644 By similarity.  
 FT CARBOHYD 76 76 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 122 122 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 220 220 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 353 353 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 459 459 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 566 566 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 723 723 S -> T (in Ref. 2).  
 SQ SEQUENCE 735 AA; 82357 MW; BE84BE115E1D0B53 CRC64;  
 Query Match 92.6%; Score 3690.5; DB 1; Length 735;  
 Best Local Similarity 91.4%; Pred. No. 1.1e-242;

```
Matches 672; Conservative 30; Mismatches 32; Indels 1; Gaps 1;
Qy 1 MW-VLLSLGLGLRMDNSFSLPQVITVPEKIRSIIEKIGIESQASYKIVIEGKPYTVNL 59
Db 1 MWRVLLSLGLGLRMDNSFSLPQVITVPEKIRSIIEKIGIESQASYKIVIEGKPYTVNL 60
Qy 60 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGLQF 119
Db 61 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGLQF 120
Qy 120 ENVSYGIEPLESSVGFPEHVIYQVKKADVSLYNEKDIESRDLISFKLQSAEPQDFAKYI 179
Db 121 ENVSYGIEPLESSVGFPEHVIYQVKKADVSLYNEKDIESRDLISFKLQSAEPQDFAKYI 180
Qy 180 EMHVIVEKQLYNHMGSDTTTVAQVFLIGITNAIFVSNITITLSLELWIDENKIATT 239
Db 181 EMHVIVEKQLYNHMGSDTTTVAQVFLIGITNAIFVSNITITLSLELWIDENKIATT 240
Qy 240 GEANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFOGKMCNDANYAGGVLLHPR 299
Db 241 GDAKELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFOGKMCNDANYAGGVLLHPR 300
Qy 300 TISLESIAVILAOQLLSMGIYDDINKKOCGSAVCIIMNPEAIFHSGVKIFSNCSFEDFA 359
Db 301 TISLESIAVILAOQLLSMGIYDDINKKOCGSAVCIIMNPEAIFHSGVKIFSNCSFEDFA 360
Qy 360 HFISKQSKQCLHNPRLDPPFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCR 419
Db 361 HFISKQSKQCLHNPRLDPPFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCR 420
Qy 420 FKAGSNCAEGPCCNCLFMSKERMCRPSFECDLPYCNCGSSASCENHYVQTGHPCGLN 479
Db 421 FKAGSNCAEGPCCNCLFMSKERMCRPSFECDLPYCNCGSSASCENHYVQTGHPCGLN 480
Qy 480 QWICIDGVCMNGDKQCDMTDFGEAEFGTECYSLNKTDSVGNCGISDGYTTOCEADNL 539
Db 481 QWICIDGVCMNGDKQCDMTDFGEAEFGTECYSLNKTDSVGNCGISDGYTTOCEADNL 540
Qy 540 QCGKLIKCYVGFLLQIPRATIIYANISGHLCIAVEFASDHADQKWKIDGTSCGSKNV 599
Db 541 QCGKLIKCYVGFLLQIPRATIIYANISGHLCIAVEFASDHADQKWKIDGTSCGSKNV 600
Qy 600 CRNQRVSSSYLVGYDCTDKNDRGVCKNKGCHCSASYLPDPCSVOSDLWPGSGISDGN 659
Db 601 CKNRQVSSSYLVGYDCTDKNDRGVCKNKGCHCSASYLPDPCSVOSDLWPGSGISDGN 660
Qy 660 PPAVAPARLPERVYNIHVSKEPWRPFFLFPFFIIFCVLIAMVKNVFORKKWRTED 719
Db 661 PPLVAVPARLPERVYNIHVSKEPWRPFFLFPFFIIFCVLIAMVKNVFORKKWRTED 720
Qy 720 YSSDEQPESESEPKG 734
Db 721 YSSDEQPESESEPKG 735

RESULT 3
Q6P2G0 PRELIMINARY; PRT; 579 AA.
AC Q6P2G0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE ADAM2 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
```

```
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC064547; AAH64547.1; -
DR InterPro; IPR006586; ADAM_Cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR Pfam; PF00200; Disintegrin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; 1.
DR PROSITE; PS00214; DISINTEGRIN_2; 1.
SQ SEQUENCE 579 AA; 64795 MW; 1F7AF0CDA2A064F9 CRC64;

Query Match 76.9%; Score 3065.5; DB 2; Length 579;
Best Local Similarity 78.5%; Pred. No. 3.1e-200;
Matches 577; Conservative 0; Mismatches 1; Indels 157; Gaps 3;
Qy 1 MW-VLLSLGLGLRMDNSFSLPQVITVPEKIRSIIEKIGIESQASYKIVIEGKPYTVNL 59
Db 1 MWRVLLSLGLGLRMDNSFSLPQVITVPEKIRSIIEKIGIESQASYKIVIEGKPYTVNL 60
Qy 60 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGLQF 119
Db 61 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGLQF 120
Qy 120 ENVSYGIEPLESSVGFPEHVIYQVKKADVSLYNEKDIESRDLISFKLQSAEPQDFAKYI 179
Db 121 ENVSYGIEPLESSVGFPEHVIYQVKKADVSLYNEKDIESRDLISFKLQSAEPQDFAKYI 171
Qy 180 EMHVIVEKQLYNHMGSDTTTVAQVFLIGITNAIFVSNITITLSLELWIDENKIATT 239
Db 172 ----- 171
Qy 240 GEANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFOGKMCNDANYAGGVLLHPR 299
Db 172 -----HPR 174
Qy 300 TISLESIAVILAOQLLSMGIYDDINKKOCGSAVCIIMNPEAIFHSGVKIFSNCSFEDFA 359
Db 175 TISLESIAVILAOQLLSMGIYDDINKKOCGSAVCIIMNPEAIFHSGVKIFSNCSFEDFA 234
Qy 360 HFISKQSKQCLHNPRLDPPFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCR 419
Db 235 HFISKQSKQCLHNPRLDPPFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCR 294
Qy 420 FKAGSNCAEGPCCNCLFMSKERMCRPSFECDLPYCNCGSSASCENHYVQTGHPCGLN 479
```

```

Db 295 FPAASNCAGGCCNCLPMSKRCRPSFECDDLPYCNCGSSASCENHYVQTGHPCLN 354
Qy 480 QWICIDGCMGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNL 539
Db 355 QWICIDGCMGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEAD-- 412
Qy 540 QCGKLIKCYVKGFLQIPRATIIYANISGHLCTIAVEFASDHADSKQKWKIDGTCGSGNKV 599
Db 413 -----GHLCTIAVEFASDHADSKQKWKIDGTCGSGNKV 444
Qy 600 CRNORCVSSSYLGYDCTTDCNDRGVCNNKXHCSCSASLYLPPDCSVOSDLWPGGSDSGN 659
Db 445 CRNORCVSSSYLGYDCTTDCNDRGVCNNKXHCSCSASLYLPPDCSVOSDLWPGGSDSGN 504
Qy 660 FPPVAIPARLPERRYENIYHSPKMRPFFLFPFFIIFCVLIAIMVKVNFQRKKWRTED 719
Db 505 FPPVAIPARLPERRYENIYHSPKMRPFFLFPFFIIFCVLIAIMVKVNFQRKKWRTED 564
Qy 720 YSSDEQPESESEPKG 734
Db 565 YSSDEQPESESEPKG 579

RESULT 4
ID AAH64547 PRELIMINARY; PRT; 579 AA.
AC AAH64547;
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE ADAM2 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore S.I., Wang J., Haie H.F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC064547; AAH64547.1; -
SQ SEQUENCE 579 AA; 64795 MW; 187AF0CDA2A064F9 CRC64;

Query Match 76.9%; Score 3065.5; DB 2; Length 579;
Best Local Similarity 78.5%; Pred. No. 3.1e-200;
Matches 577; Conservative 0; Mismatches 1; Indels 157; Gaps 3;

Qy 1 MW-VLFLSLGGLGRLMDSNFDLPVQITVPEKIRSIIEKIESQASYKIVIEGKPYTNL 59

```

```

Db 1 MWRVLFLLSLGGLGRLMDSNFDLPVQITVPEKIRSIIEKIESQASYKIVIEGKPYTNL 60
Qy 60 MQKNFLPHNFRVYSYSGTGMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQF 119
Db 61 MQKNFLPHNFRVYSYSGTGMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQF 120
Qy 120 ENVSYGIEPLESSVGFPHYIYQVHKKADVSLYNEKDIERSDLFSFKLOSABPQDFAKYI 179
Db 121 ENVSYGIEPLESSVGFPHYIYQVHKKADVSLYNEKDIERSDLFSFKLOSABPQDFAKYI 171
Qy 180 EMHVIVEKQLYNHMGSDTTVAQKVFQILGTNAIFVSPNITITLSSLELWIDENKATTT 239
Db 172 ----- 171
Qy 240 GEANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMCDANDYAGGVWLHPR 299
Db 172 -----HPR 174
Qy 300 TISLESIAVILAQLLSLMSGITYDDINKCOCGAVCINMPEATHFSGVKIFSNCSFEDFA 359
Db 175 TISLESIAVILAQLLSLMSGITYDDINKCOCGAVCINMPEATHFSGVKIFSNCSFEDFA 234
Qy 360 HFTSKQKSOCLHNQPRLDPPFKQAVCNKAKLEAGEECDCGTEQDCALIGETCCDIATCR 419
Db 235 HFTSKQKSOCLHNQPRLDPPFKQAVCNKAKLEAGEECDCGTEQDCALIGETCCDIATCR 294
Qy 420 FKAGSNCAEGPCCECNCFLMSKRCRPSFECDDLPEYCNCGSSASCENHYVQTGHPCLN 479
Db 295 FKAGSNCAEGPCCECNCFLMSKRCRPSFECDDLPEYCNCGSSASCENHYVQTGHPCLN 354
Qy 480 QWICIDGCMGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNL 539
Db 355 QWICIDGCMGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEAD-- 412
Qy 540 QCGKLIKCYVKGFLQIPRATIIYANISGHLCTIAVEFASDHADSKQKWKIDGTCGSGNKV 599
Db 413 -----GHLCTIAVEFASDHADSKQKWKIDGTCGSGNKV 444
Qy 600 CRNORCVSSSYLGYDCTTDCNDRGVCNNKXHCSCSASLYLPPDCSVOSDLWPGGSDSGN 659
Db 445 CRNORCVSSSYLGYDCTTDCNDRGVCNNKXHCSCSASLYLPPDCSVOSDLWPGGSDSGN 504
Qy 660 FPPVAIPARLPERRYENIYHSPKMRPFFLFPFFIIFCVLIAIMVKVNFQRKKWRTED 719
Db 505 FPPVAIPARLPERRYENIYHSPKMRPFFLFPFFIIFCVLIAIMVKVNFQRKKWRTED 564
Qy 720 YSSDEQPESESEPKG 734
Db 565 YSSDEQPESESEPKG 579

RESULT 5
Q866A8
ID Q866A8 PRELIMINARY; PRT; 735 AA.
AC Q866A8;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Fertilin beta.
GN Name=FTNB;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Testis;
RX MEDLINE=22872989; PubMed=14510675;
RA Day A.E., Quilter C.R., Sargent C.A., Mileham A.J.;
RT "Chromosomal mapping, sequence and transcription analysis of the
RT porcine fertilin beta gene (ADAM2).";
RL Anim. Genet. 34:375-378 (2003).

```

```

[2]
SEQUENCE FROM: N.A.
RP TISSUE-Testis;
RA Day A.; (2004), Department of Applied Sciences, Anglia Polytechnic
RL University, Cambridge, United Kingdom.
DR EMBL; AJ309003; CAC84225.2; -.
DR HSP; P18619; 1FVL.
DR MEROPS; M12.950; -.
DR GO; GO:004322; F:metalloendopeptidase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR008994; Nucleic_acid OB.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR Pfam; PF00200; Disintegrin; 1.
DR Pfam; PF01562; Pep.M12B_propep; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00215; ADAM_MERO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; 1.
DR PROSITE; PS00214; DISINTEGRIN_2; 1.
DR PROSITE; PS00026; EGF_3; 1.
SQ SEQUENCE 735 AA; 81804 MW; 16F5554690A8F4AE CRC64;

Query Match 70.8%; Score 2820; DB 2; Length 735;
Best Local Similarity 68.6%; Pred. No. 2.1e-183;
Matches 502; Conservative 96; Mismatches 132; Indels 2; Gaps 2;

QY 3 VLFLSLGLGLMDNFDSPVQITVPEKRSIIKEGIESQASVKIVIEGKPYTNLMQK 62
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4 LLFLLLGLGLQADSSERLRVQITVPEKRSISSEGVESHVSNIIIEGKTYTNLVQK 63

QY 63 NPLPHNRVSYSGTGMKLPDQFQNFCHQYVIEGPKSVVMVSTCTGLRGLQFENV 122
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
64 SFLPHNRVYGYNGTGMKPLEQFQNFCHQYVIEGPKSVVMVSTCTGLRGLQFENV 123

QY 123 SYGTEPLESSVGRHVIYQVKKADVSLYNEKDIERSDLFSKLQSRPQDPAKYTEMH 182
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
124 SYGTEPLESPVGRHVIYQVKKADVSLYNEKDIERSDLFSKLQSRPQDPAKYTEMH 183

QY 183 VIVEKQLYNHMGSDTTVVAQVFLQIGLTNAIFVSFNITILSLLELWIDENKIATTGEA 242
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 IVEKQLYNHMGSDTAVVQKIFQIGLTNAIFTSFNITILSLLELWIDENKITVTGDA 243

QY 243 NELLHTEFLRWKTSVLVLRPHDVAFLVYREKSNVVGATFQGMCDANYAGVGLVHPHTIS 302
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
244 NELLHTEFLRWKTSVLVLRPHDVAFLVYREKSNVVGATFQGMCDANYAGVGLVHPHTIT 303

QY 303 LESLAVILAQLLSLMSGTYDDINKKCCGAGVATMNPFAHFGSKVIFSNCSPEDEFAHI 362
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 LESLAVTMAQLLSLMSGTYDDINKKCCGAGVATMNPFAHFGSKVIFSNCSPEDEFAHFT 363

QY 363 SKQSKQLCHNPRLDPPFFKQAVCGNAKLEAGEBCDGTEDCALIGETCCDIATCRFKA 422
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
364 AKSKQSKQLCHNPRLDPPFFKQAVCGNAKLEAGEBCDGTEDCALIGETCCDIATCRFKA 422

QY 423 GSNCAEGGCCNCLFMSKERNCRSPFBCDLPFCVNGSSASCPENHYVQTHCPGLNQWI 482
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
423 GSACATGPGCCDSCFMKAGQTCRLTDECDLLEYCNGSSAACQBDLYVQDGHPCSDNQWL 482

QY 483 CIDGVCMSGDKQCTDTTFQKEVFPFSECYSHLNSKTDVSGNCGISDSGYTTCEDADNLQCG 542
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
483 CVQKCKISGMKQCSSETFGDGNANYGTEDCFTHLNSKTDVSGNCGISDSGYTTCEDADNLQCG 542
```



DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR008994; Nucleic acid OB.  
 DR InterPro; IPR001590; Peptidase M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR Pfam; PF00200; Disintegrin; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR PRINTS; PR00289; DISINTEGRIN.  
 DR ProDom; PD000664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS00215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; 1.  
 DR PROSITE; PS00214; DISINTEGRIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 KW Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
 FT SIGNAL 1 16 Potential.  
 FT PROPEP 17 174 By similarity.  
 FT CHAIN 175 745 ADAM 2.  
 FT DOMAIN 17 696 Extracellular (Potential).  
 FT TRANSMEM 697 717 Potential.  
 FT DOMAIN 718 745 Cytoplasmic (Potential).  
 FT DOMAIN 175 382 Metalloprotease-like.  
 FT DOMAIN 383 474 Disintegrin-like.  
 FT DOMAIN 475 605 Cys-rich.  
 FT DOMAIN 611 644 EGF-like.  
 FT DOMAIN 148 151 Poly-Ser.  
 FT DISULFID 287 370 By similarity.  
 FT DISULFID 329 354 By similarity.  
 FT DISULFID 444 457 Potential.  
 FT DISULFID 615 626 By similarity.  
 FT DISULFID 620 632 By similarity.  
 FT DISULFID 634 643 By similarity.  
 FT CARBOHYD 122 122 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 220 220 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 353 353 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 458 458 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 558 558 N-linked (GLNAC. .) (Potential).  
 SQ SEQUENCE 745 AA; 83150 MW; B5D8DC0168999800 CRC64;  
 Query Match 68.3%; Score 2723; DB 1; Length 745;  
 Best Local Similarity 64.8%; Pred. No. 8.4e-177;  
 Matches 481; Conservative 103; Mismatches 146; Indels 12; Gaps 2;  
 QY 3 VLFLLGLGLRMDNSFLVQITVPEKIRSIKEGIESQASYKIIEGKPYTVNLMQK 62  
 DB 4 LFLLLGLTGLQDDNSERLRVQTVPEKIRSTSGGVETHVSYILLIEGKTYTVNLMQK 63  
 QY 63 NFLPHFRVSYSGTGIMKPLDQFQNFCHYQYIEGPKSVVMVSTCTGLRGVLQFENV 122  
 DB 64 AFLPHFRVSYSGTGIMKPLEHEFQNFCHYQYIEGPKSVMAISTCTGLRGVLQFENV 123  
 QY 123 SYGIEPLESSVGPEHVIYQVKKKADVSLNEXDKIESRDLSPKLSQAEPOODFAKVIEMH 182  
 DB 124 SYGIEPLESGPEHVIYQVKKKADVSLNEXDKIESRDLSPKLSQAEPOODFAKVIEMH 183  
 QY 183 VIVEKOLYNHMGSDTTVAQVFLQIGLITNAIVFSFNITILSLSLWIDENKIATGEA 242  
 DB 184 IVEKOLYNHMGSDTTVAQVFLQIGLITNAIVFSFNITILSLSLWIDENKIPTGDA 243  
 QY 243 NELLHTFLRWKTSYVLRLPHDVAFLVLYREKSNVYGATFGKWCNDANYAGVVLHRTIS 302  
 DB 244 NELLHFRVWKRYSYVLRLPHDVAFLVLYREKSNVYGATFGKWCNDANYAGVVLHRTIS 303  
 QY 303 LESLAVILAQLLSLWIGITVDDINKKCCSACVIMNPEAHFSGVKIFSNCSDEFAHFI 362  
 DB 304 LESLAVILAQLLSLWIGITVDDINKKCCSACVIMNPEAHFSGVKIFSNCSDEFAHFI 363  
 QY 363 SKOKSOCLHNQPRLDPPFFKQQA VCGNAKLEAGBECDCGTEDDCALIGETCCDIATCRFKA 422  
 DB 364 SKPKSOCLHNQPRLDPPFFKQQA VCGNAKLEAGBECDCGTEDDCALIGETCCDIATCRFKA 422  
 QY 423 GSNCAEGPCCECLFWSKERMCRPSPEECDLPYCVNGSSACBPENHYVOTGHPGCLNQWI 482  
 DB 423 GSACDTGLCCESCAFIPKGHICRGSTDECDLHYICNGSSAACQEDVVYVQDHPGCGNQWL 482  
 QY 483 CIDGVCMGSGDKQCTDTFGKEVEFGPSECHSLNASKTDSVNGCGISDSGYTQCEADNLQCG 542  
 DB 483 CIGSICVDGIGKQCFDIFGEGTSYAPAECPORLNSMNDLSCNGCVTPGTPCTSENVRG 542  
 QY 543 KLIKCVKGLLOIPRATIIYANISGHLCTIAVFASDHADSQKWIKDGTSKCSNKKVCRN 602  
 DB 543 KLLCTYDKREVISVENASVMYNSINGKICIGLHYEGNEDEGMMVWKDGTVCGESKICQN 602  
 QY 603 QRCVSSSYLGVDCTTDKCDNRGVCKNKKCHCSASYLPDSCSVQSDLWPGSGSDSGNFP 662  
 DB 603 KQCVSSFLNYDCNPKCNQGVCKNKKCHCSASYLPDSCSVQSDLWPGSGSDSGNFP 662  
 QY 663 VAIP-----ARLPERRYIENIYHSKPMRWPFFLFPFFIIPFCVLIAIMVKVNFQ 711  
 DB 663 SEPTGGAFTDVGTTPLAESRYIENYRSKPTRWPFLLFPFFIIPFCVLIAIMVKVNFQ 722  
 QY 712 RKKWRTEDYSSDQPESESEPK 733  
 DB 723 RKKWRTEDYSSDQPESESEPK 744  
 RESULT 7  
 AD02\_RABIT STANDARD; PRT; 751 AA.  
 AC Q28660;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE ADAM 2 precursor (A disintegrin and metalloprotease domain 2)  
 DE (Fertilin beta subunit) (PH-30) (PH30) (PH30-beta).  
 GN Name=ADAM2; Synonyms=FTNB;  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97071141; PubMed=8914066;  
 RA Hardy C.M., Holland M.K.;  
 RT "Cloning and expression of recombinant rabbit fertilin.";  
 RL Mol. Reprod. Dev. 45:107-116(1996).  
 CC -|- FUNCTION: Sperm surface membrane protein that may be involved in  
 CC sperm-egg plasma membrane adhesion and fusion during  
 CC fertilization. Could have a direct role in sperm-zona binding or  
 CC migration of sperm from the uterus into the oviduct. Interactions  
 CC with egg membrane could be mediated via binding receptors on the  
 CC disintegrin-like domain to one or more integrins receptors on the  
 CC egg. This is a non catalytic metalloprotease-like protein (By  
 CC similarity).  
 CC -|- SUBUNIT: Heterodimer with ADAM1/fertilin alpha.  
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -|- TISSUE SPECIFICITY: Expressed specifically in testis.  
 CC -|- DOMAIN: A tripeptide motif (VGE) within disintegrin-like domain  
 CC could be involved in the binding to egg integrin receptor and thus  
 CC could mediate sperm/egg binding (By similarity).  
 CC -|- PTM: The signal and the metalloprotease domain are cleaved during  
 CC the epididymal maturation of the spermatozoa (By similarity).  
 CC -|- SIMILARITY: Belongs to peptidase family M12B.  
 CC -|- SIMILARITY: Contains 1 disintegrin domain.  
 CC -|- SIMILARITY: Contains 1 EGF-like domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC  EMBL; U46070; AAA93321.1; -.
CC  HSSP; P30403; 1N4Y.
CC  MEROPS; M12.950; -.
CC  InterPro; IPR006586; ADAM cysteine.
CC  InterPro; IPR001762; Disintegrin.
CC  InterPro; IPR000742; EGF_2.
CC  InterPro; IPR006209; EGF-like.
CC  InterPro; IPR008994; Nucleic acid OB.
CC  InterPro; IPR001590; Peptidase M12B.
CC  InterPro; IPR002870; Peptidase M12B_N.
CC  Pfam; PF00200; Disintegrin; 1.
CC  Pfam; PF01562; Pep_M12B_propep; 1.
CC  Pfam; PF01421; Reprolysin; 1.
CC  PRINTS; PD00289; DISINTEGRIN.
CC  ProDom; PD000664; Disintegrin; 1.
CC  SMART; SM00608; ACR; 1.
CC  SMART; SM00050; DISIN; 1.
CC  PROSITE; PS50215; ADAM_MEPRO; 1.
CC  PROSITE; PS00427; DISINTEGRIN_1; 1.
CC  PROSITE; PS50214; DISINTEGRIN_2; 1.
CC  PROSITE; PS00022; EGF_1; FALSE NEG.
CC  PROSITE; PS01186; EGF_2; FALSE NEG.
CC  PROSITE; PS50026; EGF_3; 1.
KW  Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.
FT  SIGNAL 16 Potential.
FT  PROPEP 17 173 By similarity.
FT  CHAIN 17 751 ADAM 2.
FT  DOMAIN 17 702 Extracellular (Potential).
FT  TRANSMEM 703 723 Potential.
FT  DOMAIN 724 751 Cytoplasmic (Potential).
FT  DOMAIN 174 381 Disintegrin-like.
FT  DOMAIN 382 474 Cys-rich.
FT  DOMAIN 475 606 EGF-like.
FT  DOMAIN 612 645 By similarity.
FT  DISULFID 286 369 By similarity.
FT  DISULFID 328 353 Potential.
FT  DISULFID 444 457 Potential.
FT  DISULFID 616 627 By similarity.
FT  DISULFID 621 633 By similarity.
FT  DISULFID 635 644 By similarity.
FT  CARBOHYD 122 122 N-linked (GlcNAc... ) (Potential).
FT  CARBOHYD 147 147 N-linked (GlcNAc... ) (Potential).
FT  CARBOHYD 219 219 N-linked (GlcNAc... ) (Potential).
FT  CARBOHYD 289 289 N-linked (GlcNAc... ) (Potential).
FT  CARBOHYD 352 352 N-linked (GlcNAc... ) (Potential).
FT  CARBOHYD 434 434 N-linked (GlcNAc... ) (Potential).
FT  CARBOHYD 458 458 N-linked (GlcNAc... ) (Potential).
FT  CARBOHYD 559 559 N-linked (GlcNAc... ) (Potential).
FT  CARBOHYD 566 566 N-linked (GlcNAc... ) (Potential).
FT  CARBOHYD 643 643 N-linked (GlcNAc... ) (Potential).
SQ  SEQUENCE 751 AA; 83616 MW; 824C6FBFC4A9FCE1 CRC64;

Query Match 67.4%; Score 2685; DB 1; Length 751;
Best Local Similarity 64.4%; Pred. No. 3.3e-174;
Matches 482; Conservative 106; Mismatches 142; Indels 18; Gaps 4;

QY 3 VLFLSLGLGRMDSNFDSPVQITVPEKIRSIIEKIGESQASYKIVIEGKPYTVNLMMQK 62
DB 4 VLFLSLGLSLGRKENSERLHVQVTVPEKMSVTSSEGFIEVYVNIIEGKTYTLNLMQK 63

QY 63 NPLPHFRVYSYSGTGIMKPLDQDFQNFCHQYVIEGPKSVVMVSTCTGLRGLQFENV 122
DB 64 LFLPRDRVRYGDSGTIMKDLQEQFQNFYQYQYQYQYQYQYQYQYQYQYQYQYQY 123

QY 123 SYGIEPLESSVGFPHVYQVKKKADVSLYNEKDIESRDLSPKLSQAEPOQDFAKYIEMH 182
DB 124 TYGIEPLESSIGFPHVYQVKKKESISLYAEKEVEFRDLPYKQSVQV-REFSQYIEMH 182

```

```

QY 183 VIVEKOLYNHMGSDTTVAQKVFQQLIGLTNAIVSFNFIITILSSLELWIDENKIATGCEA 242
DB 183 VVEKNLYKMGSDTAVVSKIFQQLIGLTNAVTSFNFIITILSSLELWIDENKISTTGDA 242

QY 243 NELLHTPLRWKTSYLVLRPHDVAFLVYREKSNVYVGTATFOGKMCDDANYAGVVLHPKTI 302
DB 243 NELLYRFLKWKSGYLVLRPHDVAFLVYREIAKYVGTATFOGKMCDDINYSGGVALHPKTI 302

QY 303 LESLAVILAOLLSISMGITVDDINKKOCGSAVCIIMNPEAHFSGVKTFNSCSPEDEFAHFI 362
DB 303 LESLAVILAOLLSISMGIAVDDINKKCPGSCVICIMNPEAHFSSGVKTFNSCSIEDFSRFI 362

QY 363 SKOKSQCLHNPRLDPPFFKQAVCGNAKLEAGBECDCGTQDQDCALIGETCCDIATCRFKA 422
DB 363 SKOKSQCLQNLPRLEPFYKQDAVCGNSIVEAGEMCDGTAECGRAVPICCNSATCRLEV 422

QY 423 GSNCAEGPCCECLFMSKERMCRPSFEEDLPYCNCGSSASCENHYVQTHGPCGLNQWI 482
DB 423 GSQCAEGCECCENTCFKRGQSCRPPVGECDLFFCYNGTSALCPDDIVIQNGHPCGENQWI 482

QY 483 CIDGVCMSGDKQCTDTFGKEVEFGPSECYSHLNSKTD-VSGNGGIGSDSGVTOCEADNLQC 541
DB 483 CVDGSCISPPPEQCKSIPEBEVGGTPECYBELNAMADIILIGITHDGYKKCESGNRKC 542

QY 542 GKLICKYVVGKFLQIPRATIIYANISGHLCAIAVEFASDHADSQKMWIKDGTSCGSKNVC 601
DB 543 GKLCKHHTTNIIDIKNATIIYANVSSICLSLEYKPDHLDAKMWVNGAVCGSNKICR 602

QY 602 NQRCVSSSYGYDCTTDKCNDRGVNKKKHCHCSASYLPDPCSVQSDLPWGGSIDGNP- 660
DB 603 DKACVETVYVNLGCTLQNCNNGICNSLQHCNCNPTFLPENCISAVDERWAGGSDVSGNFQ 662

QY 661 -----PVAIPARLPERRYENIYH-SKPWEWPFFLFTPRPIIFCVLIAM 705
DB 663 GGVGPIGGLTSVGTPTFPGVGVRRYEMDVIQTAKTRWPFLIPFLIPFLLIALL 722

QY 706 VKNVORKKWRTEYSSDEQSESESESPK 733
DB 723 VKVQFQKWKTEYTSDEQFESDELK 750

RESULT 8
AD02 RAT STANDARD; PRT; 737 AA.
AC Q63202;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)
GN (Fertilin beta subunit) (PH-30) (PH30) (PH30-beta).
OS Names=Adam2; Synonyms=Ftnb;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=Sprague-Dawley; TISSUE=Testis;
RX MEDLINE=98019039; PubMed=9358007;
RA McLaughlin E.A., Frayne J., Barker H.L., Jury J.A., Jones R.,
RA Ford W.C.L., Hall L.;
RT "Cloning and sequence analysis of rat fertilin alpha and beta -
RL Mol. Hum. Reprod. 3:801-809(1997).
CC -!- FUNCTION: Sperm surface membrane protein that may be involved in
CC sperm-egg plasma membrane adhesion and fusion during
CC fertilization. Could have a direct role in sperm-zona binding or
CC migration of sperm from the uterus into the oviduct. Interactions
CC with egg membrane could be mediated via binding between its
CC disintegrin-like domain to one or more integrins receptors on the
CC egg. This is a non catalytic metalloprotease-like protein (By
CC similarity).
CC -!- SUBUNIT: Heterodimer with ADAM1/fertilin alpha.

```



egg integrins and the mouse sperm homologue of PH-30 (fertilin) beta.";  
[2]  
[3]

SEQUENCE FROM N.A.  
RA Gupta S.K., Alves K., Palladino L.O., Mark G.E., Hollis G.F.;  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

[3]  
SEQUENCE FROM N.A.  
RC STRAIN=CS7BL/6J; TISSUE=Testis;  
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nakaio I., Osato N., Saio R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriber L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Konai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Kanagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Meglott D.R., Maitais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlesstedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Barney E., Hayaishi Y.,  
RA "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
[4]

SEQUENCE OF 17-735 FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=95269891; PubMed=7750654;  
RA Wolfberg T.G., Straight P.D., Garena R.L., Huovila A.-P.,  
RA Primakoff P., Myles D.G., White J.M.;  
RT "ADAM, a widely distributed and developmentally regulated gene family  
RT encoding membrane proteins with a disintegrin and metalloprotease  
RT domain.";  
RL Dev. Biol. 169:378-383 (1995).

CC -1- FUNCTION: Sperm surface membrane protein that may be involved in  
CC sperm-egg plasma membrane adhesion and fusion during  
CC fertilization. Could have a direct role in sperm-zona binding or  
CC migration of sperm from the uterus into the oviduct. Interactions  
CC with egg membrane could be mediated via binding between its  
CC disintegrin-like domain to one or more integrins receptors on the  
CC egg. This is a non catalytic metalloprotease-like protein (By  
CC similarity).  
CC -1- SUBUNIT: Heterodimer with ADAM1/fertilin alpha.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: Expressed specifically in testis.  
CC -1- DOMAIN: A tripeptide motif (QDE) within disintegrin-like domain  
CC could be involved in the binding to egg integrin receptor and thus  
CC could mediate sperm/egg binding (By similarity).  
CC -1- PTM: The signal and the metalloprotease domain are cleaved during  
CC the epididymal maturation of the spermatozoa.  
CC -1- SIMILARITY: Belongs to peptidase family M12B.  
CC -1- SIMILARITY: Contains 1 disintegrin domain.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; U16242; AAA90980.1; -;  
CC EMBL; U38806; AAD04207.1; -;  
CC EMBL; AK016550; BAB30298.1; -;  
CC EMBL; U22057; AAA74921.1; -;  
CC HSSP; P17347; 2ECH.  
CC MEROPS; M12.950; -;  
CC MGD; MGI:1340894; Adam2.  
CC InterPro; IPR006586; ADAM cysteine.  
CC InterPro; IPR001762; Disintegrin.  
CC InterPro; IPR000742; EGF 2.  
CC InterPro; IPR006209; EGF-like.  
CC InterPro; IPR008994; Nucleic acid OB.  
CC InterPro; IPR001590; Peptidase\_M12B.  
CC Pfam; PF00200; Disintegrin; 1.  
CC Pfam; PF01562; Pep\_M12B\_prosep; 1.  
CC Pfam; PF01421; Repolysin; 1.  
CC PRINTS; PR00289; DISINTEGRIN.  
CC PRODOM; PD000664; Disintegrin; 1.  
CC SMART; SM00608; ACR; 1.  
CC SMART; SM00050; DISIN; 1.  
CC PROSITE; PS00215; ADAM\_MEPRO; 1.  
CC PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
CC PROSITE; PS00214; DISINTEGRIN\_2; 1.  
CC PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
CC PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
CC PROSITE; PS00026; EGF\_3; 1.  
CC Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
FT SIGNAL 1 18 Potential  
FT PROPEP 19 180 By similarity.  
FT CHAIN 181 735 ADAM 2.  
FT DOMAIN 19 686 Extracellular (Potential).  
FT TRANSMEM 687 707 Potential.  
FT DOMAIN 708 735 Cytoplasmic (Potential).  
FT DOMAIN 181 388 Metalloprotease-like.  
FT DOMAIN 389 479 Disintegrin-like.  
FT DOMAIN 480 611 Cys-rich.  
FT DOMAIN 615 648 EGF-like.  
FT DISULFID 293 376 By similarity.  
FT DISULFID 335 360 By similarity.  
FT DISULFID 449 462 Potential.  
FT DISULFID 619 630 By similarity.  
FT DISULFID 624 636 By similarity.  
FT DISULFID 638 647 By similarity.  
FT CARBOHYD 128 128 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 226 226 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 279 279 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 359 359 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 463 463 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 489 489 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 569 569 N-linked (GLCNAC. .) (Potential).  
FT CARBOHYD 585 585 N-linked (GLCNAC. .) (Potential).  
FT CONFLICT 2 2 W -> R (in Ref. 2).  
FT CONFLICT 17 20 LSQS -> IRHE (in Ref. 4).  
FT CONFLICT 24 25 GT -> M (in Ref. 4).  
FT CONFLICT 113 113 I -> M (in Ref. 1).  
FT CONFLICT 234 242 LEFWMENK -> WNFQWKKQ (in Ref. 4).  
FT CONFLICT 246 247 TG -> QA (in Ref. 4).  
FT CONFLICT 286 286 A -> L (in Ref. 4).  
FT CONFLICT 331 332 DV -> RRL (in Ref. 4).  
FT CONFLICT 382 382 R -> T (in Ref. 2).  
FT CONFLICT 658 658 S -> T (in Ref. 4).  
FT CONFLICT 679 679 A -> R (in Ref. 2).  
FT CONFLICT 712 712 Q -> P (in Ref. 1).  
SQ SEQUENCE 735 AA; 82374 MW; 75EC8523CF5B8E2B CRC64;

Query Match 61.6%; Score 2453.5; DB 1; Length 735;  
Best Local Similarity 59.7%; Pred. No. 1.8e-158;



FT	TRANSMEM	681	701	Potential.	AD18_MACFA	STANDARD;	PRT;	746 AA.
FT	DOMAIN	702	735	Cytoplasmic (Potential).	Q95134;			
FT	DOMAIN	178	382	Metalloprotease-like.	28-FEB-2003 (Rel. 41, Created)			
FT	DOMAIN	383	472	Disintegrin-like.	28-FEB-2003 (Rel. 41, Last sequence update)			
FT	DOMAIN	473	609	Cys-rich.	01-OCT-2004 (Rel. 45, Last annotation update)			
FT	DOMAIN	610	643	EGF-like.	ADAM 18 precursor (A disintegrin and metalloproteinase domain 18)			
FT	DISULFID	287	370	By similarity.	(Transmembrane metalloproteinase-like, disintegrin-like, and cysteine-rich protein III) (TMDC III).			
FT	DISULFID	329	354	By similarity.	Name=ADAM18; Synonyms=TMDC3;			
FT	DISULFID	442	455	Potential.	Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).			
FT	DISULFID	614	625	By similarity.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
FT	DISULFID	619	631	By similarity.	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;			
FT	DISULFID	633	642	By similarity.	Cercopithecinae; Macaca.			
SQ	SEQUENCE	735 AA;	81904 MW;	7535FC39F44FB645 CRC64;	NCBI_TaxID=9541;			
Query Match					SEQUENCE FROM N.A.			
Best Local Similarity					TISSUE=Testis;			
Matches					MEDLINE=98328566; PubMed=9665629;			
Qy	3	VLFLSLGLGL--RMSDNFSLPVOITVPEKIRSIKEGIESQASYKIVIEGKPYTVNLM	60		Frayne J., Jury J.A., Barker H.L., Perry A.C.F., Jones R., Hall L.;			
Db	4	LLLLLCLASLGLGPKKYYVNSGQITVPEKIRSVKGSVESEVYKIVIENTYIIVNLV	63		"Macaque MDC family of proteins: sequence analysis, tissue			
Qy	61	QKPLPHNPFVYSGTGIMKPLDQDQNFCHYQGYIEGPKSVVMVSTCTGLRGVLOFE	120		RT distribution and processing in the male reproductive tract.";			
Db	64	RKMFLPHDFQVYSYDSAGIMKPFEDYSQFCYQGHIEGPTSLASITCAGLRGLLOFE	123		Mol. Hum. Reprod. 4:429-437(1998).			
Qy	121	NVSGIEPLSSGVFEHVIVQVHKHADVSLYNEKDIESRDLSPKLOSAPQODFAKYIE	180		-1- FUNCTION: Sperm surface membrane protein that may be involved in			
Db	124	TVSGIEPLKSSIGFEHVIVPKHDN-EKSYLKKSINVKVYKISKSVR-THYIE	181		spermatogenesis and fertilization. This is a non catalytic			
Qy	181	MHVIVEKOLYNHMGSDTTVAQVQFOLIGLTNAIFVSNFTIILSSLELWIDENKIATTG	240		metallopotease-like protein (By similarity).			
Db	182	MHIIVEKNLYKMGNTATVTEKIFQLVGLMNAFFSLNLTIVLASLELWIDENKIPTTG	241		-1- SUBCELLULAR LOCATION: Type I membrane protein.			
Qy	241	EANELHTFLRWKTSYLVLPDVAFLVYREKSNVYGATFOGKCDKANYAGGVVLPRT	300		-1- TISSUE SPECIFICITY: Expressed predominantly in adult and			
Db	242	DVNELLHAFQWKRSYLVLPDVAFLVYRESPTIYGAIFQGMICNTSVGGGIALHSKT	301		prepubertal testis.			
Qy	301	ISLESVALIAQLLSMGITDIDINKCQSGAVCIMNPAIHPSGVKIFSNCSFEDFAH	360		-1- DOMAIN: A tripeptide motif (ECD) within disintegrin-like domain			
Db	302	ITLDSFGVILLVSLVSGMAYDNADLCRCGALCLMSPEAVFSSGMKMFNSCSVKAFKL	361		could be involved in the binding to egg integrin receptor and thus			
Qy	361	FISKQKSQCLHNPRLDPFFKQQAQVGNAKLEAGECCDGTGEOCALIGTCCDIATCRF	420		-1- PTM: The prodomain and the metalloprotease-like domain are cleaved			
Db	362	FTSSQVSQCLNQPYLFPVYRSNPNVCGNNRVEQGEDCDGSEQCQ---DTCCDAATCRL	418		during the epididymal maturation of the spermatozoa.			
Qy	421	KAGSNCAEGCCENCLFWSKERMCRPSFECDLPEYCNCGSSASCPENHYVQTGHPCLNQ	480		-1- SIMILARITY: Belongs to peptidase family M12B.			
Db	419	KSTSRCAQGCCNCEPKTKEVCRESDECDLPEYCNCGSGACQEDLYVINGHRCAEE	478		-1- SIMILARITY: Contains 1 disintegrin domain.			
Qy	481	WICIDGVCMSGDKQCTDTFGKEVFGPSECVSHLSKNTDVSNGCI--SDSGYTTQCEADNL	539		-1- SIMILARITY: Contains 1 EGF-like domain.			
Db	479	WICNNGRCLSGKAQVQTFGTGEMSGVDCFEQLNTKNDITGNGCILLSPGNYKACASNW	538		This SWISS-PROT entry is copyright. It is produced through a collaboration			
Qy	540	QCGKLIKYVGKFLLOIPRATIIYANTISGHLCIAVEFASHDASQKWKIKDGTSCGNKV	599		between the Swiss Institute of Bioinformatics and the EMBL outstation -			
Db	539	KCGKLIYSKSETLRNEKGTIIYANTISGHICVSEIYPPGHAKSALMWVRDGTGCGPSEV	598		the European Bioinformatics Institute. There are no restrictions on its			
Qy	600	CRNQRVSSSVLYGDCDTCNDKNDGVCNKKHCHCSASLYLPDSCVSDLWPGSSIDSGN	659		use by non-profit institutions as long as its content is in no way			
Db	599	CRQOCCVSSSVLYGDCDTPATCSDRHGVNKKRKHCHCNPTYPNCEQDSTKPGSVSDSGN	658		modified and this statement is not removed. Usage by and for commercial			
Qy	660	--PPVAIPARLPERRVIENTIYHSPKRWPPFLFIPFFIIPCVLIAMVKVNFORKKWT	717		entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
Db	659	LYEPI-----PTYFVEGAYHTKSKWPPFLIIPFVIFSVLVAIVKVVYQKKWKT	712		or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
Qy	718	EDVSDSDQPSSESPK 733			EMBL; Y08617; CAA69909.1; --			
Db	713	EDVANDENIESESPK 728			HSP; P83469; IMPZ.			
RESULT 11					MEROFS; M12.957; --			
AD18_MACFA					InterPro; IPR006586; ADAM_cysteine.			
					InterPro; IPR001762; Disintegrin.			
					InterPro; IPR000742; EGF_2.			
					InterPro; IPR006209; EGF-like.			
					InterPro; IPR001590; Peptidase_M12B.			
					InterPro; IPR002870; Peptidase_M12B_N.			
					Pfam; PF00200; Disintegrin; 1.			
					Pfam; PF01562; Pep_M12B_propep; 1.			
					Pfam; PF01421; Reprolysin; 1.			
					ProDom; PD000664; Disintegrin; 1.			
					SMART; SM00608; ACR; 1.			
					SMART; SM00050; DISIN; 1.			
					PROSITE; PS00215; ADAM_MERO; 1.			
					PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.			
					PROSITE; PS00214; DISINTEGRIN_2; 1.			
					PROSITE; PS00022; EGF_1; FALSE_NEG.			
					PROSITE; PS01186; EGF_2; FALSE_NEG.			
					PROSITE; PS00026; EGF_3; 1.			
					EGF-like domain; Glycoprotein; Signal; Spermatogenesis; Transmembrane.			
					FT SIGNAL	1	16	Potential.
					FT PROPEP	17	183	Potential.
					FT CHAIN	184	746	ADAM 18.
					FT DOMAIN	177	687	Extracellular (Potential).
					FT TRANSMEM	688	708	Potential.
					FT DOMAIN	709	746	Cytoplasmic (Potential).

Query Match 39.8%; Score 1586.5; DB 1; Length 746;  
 Best Local Similarity 42.4%; Pred. No. 1.9e-99;  
 Matches 310; Conservative 125; Mismatches 261; Indels 35; Gaps 12;

FT DOMAIN 184 389 Metalloprotease-like.  
 FT DOMAIN 390 479 Disintegrin-like.  
 FT DOMAIN 480 619 Cys-rich.  
 FT DOMAIN 620 654 EGF-like.  
 FT DISULFID 293 376 By similarity.  
 FT DISULFID 335 360 By similarity.  
 FT DISULFID 458 483 Potential.  
 FT DISULFID 624 636 By similarity.  
 FT DISULFID 630 642 By similarity.  
 FT DISULFID 644 653 By similarity.  
 FT CARBOHYD 36 36 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 122 122 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 149 149 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 156 156 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 177 177 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 294 294 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 359 359 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 465 465 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 611 611 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 625 625 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 746 AA; 84448 MW; 1D8C7E9071502E30 CRC64;

39.8%; Score 1586.5; DB 1; Length 746;  
 42.4%; Pred. No. 1.9e-99;  
 125; Conservative  
 261; Mismatches  
 35; Indels  
 12; Gaps

Qy 2 WVLFLSLGLGRMDSNFDLSLPQVITPEKIRSIIRKIEGIESOASQYKIVIEGKPYTVNLQM 61  
 Db 3 FLALLTELGRQAHVSGEGFLHVTVPKILSNDSSEVSRKMIYITIDGQPYTLHLRK 62

Qy 62 KNFLPHNFRVSYSGTGMKPLDQDFQNFCHYQYIEGYPKSVVMVSTCTGLRGVLQFEN 121  
 Db 63 QSELPQNFVLYTYNEAGSLHSESPYFMHCHYQYAAEFNPSFVTLSCSLGRGLQFEN 122

Qy 122 VSGIEPLESSVGEFHVIVYQKKADVSLY--NEKDIESRDLSPKLOSAEPQODPAK-- 177  
 Db 123 VSGIEPLESSARFEHIIYQMKNDPNVSIHAENYSHIWQDOQSYKVPILNSQKKNKSKLL 182

Qy 178 --YIEMHVIKQLYNHMGSDTTVAQVFLQGLTNAIFVSNFIIITLSLELWIDENK 235  
 Db 183 PQYLEIIVIEKALDYMGSEMAVTKIVQVIGLVNMTQFLTLVTLSSLELWSNENQ 242

Qy 236 IATTGEANELLHLFLRWKTSYLVLRPHDVAFLVYREKSNVYGATPGKMKCDANYAGGVV 295  
 Db 243 ISTSGDADDLQRLAWKRDYLLRPHDIAYLLVYRKHYPYVGATFGTTCNESYDAGIA 302

Qy 296 LHPTISLESIAVLAQLLSLWGITVDDINKOCSCGAVCIMMPEATHFSGVIFSNCSF 355  
 Db 303 MYPDALDEGFSVILIAQLLGLNVLTYDDITQCFCLRATCIMNHEAMARGIKIFNSCM 362

Qy 356 EDFAHFLSKQSOCLHNQPRLDPPFFKQQAQVGNKALBAGBECDCGTQDCALIGETCCDI 415  
 Db 363 HDYRYFVSKFEAKLQKLSNLQPLHONQVCGNGILSNBECDCGNKKECQF--KKCCDY 420

Qy 416 ATCRFKAGSNACGPPCC--ENCLFMSKRMCRPFE--ECDLPFVYNGSSASCENHYVQTG 473  
 Db 421 NTKLKGSVKCGSPCCTSKLSIVGTPCRKSVDPCECTEYCNGTSGSDCPVPTVALNG 480

Qy 474 HPCLGLNQWICIDGVMSGDQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQ 533  
 Db 481 HLCKLGTAYCYNGQCQTTDNOCAKIFGKAQGAFFACFKVNSLHETSENGCFKNSQPLP 540

Qy 534 CEADNLQCGKLI CKYVGKFLQIPRATIIIVANTSGHLCAIV---EPASDHADSQKWK 589  
 Db 541 CERKDVLCGLKACVQPHNAYKSDIQYTVSYIQDVCHVCIATGSSNRSRSGTDN--AYVA 598

Qy 590 DGTSCGNKVCNRCNORCVSSSYLGYDC--TTDKCNDRGVCKNKKCHCSASYLPDPCVSQSD 648  
 Db 599 DGTMGCPMEYCVNKTCKRVHLTGVCNCTTKCKGKGCNNFGNCGQCPGPKHPDCKPQFG 658

Qy 649 LWFQGSIDSNGFPFVAPLPRERYIENYHSPMEW---PPFLFPFPFIIFCVLIAM 705  
 Db 659 S-PGGSIDGDNF-----QKSDEFYTEKGYNAHNNWNFVTLFVILVFPFIITIVI--- 707

Qy 706 VKVNFQKKWR 716  
 Db 708 ----FKRNEIR 714

RESULT 12  
 AD18 HUMAN  
 ID AD18\_HUMAN STANDARD; PRT; 739 AA.  
 AC Q9Y307;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE ADAM 18 precursor (A disintegrin and metalloproteinase domain 18)  
 DE (Transmembrane metalloproteinase-like, disintegrin-like, and cysteine-rich protein III) (tMDC III).  
 GN Name=ADAM18; Synonyms=TMDC3;  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA Hall L., Frayne J., Dimsey E.A.;  
 RT "Nucleotide sequence of the human tMDC III sperm surface protein  
 Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RL Submittted";  
 CC -I- FUNCTION: Sperm surface membrane protein that may be involved in  
 spermatogenesis and fertilization. This is a non catalytic  
 metalloproteinase-like protein (By similarity).  
 CC -I- SUBCELLULAR LOCATION: Type I Membrane protein.  
 CC -I- TISSUE SPECIFICITY: Expressed specifically in testis.  
 CC -I- DOMAIN: A tripeptide motif (ECD) within disintegrin-like domain  
 could be involved in the binding to egg integrin receptor and thus  
 could mediate sperm/egg binding (By similarity).  
 CC -I- PTM: The prodomain and the metalloproteinase-like domain are cleaved  
 during the epididymal maturation of the spermatozoa (By  
 similarity).  
 CC -I- SIMILARITY: Belongs to peptidase family M12B.  
 CC -I- SIMILARITY: Contains 1 disintegrin domain.  
 CC -I- SIMILARITY: Contains 1 EGF-like domain.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>  
 or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 -----

CC AJ133004; CAB40812.1; --  
 DR HSP; P83469; IMPZ.  
 DR MEROPS; M12.957; --  
 DR Genew; HGNC:196; ADAM18.  
 DR GO; GO:0005624; C:membrane fraction; TAS.  
 DR GO; GO:0008237; F:metalloproteinase activity; TAS.  
 DR GO; GO:0007283; P:spermatogenesis; TAS.  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR Pfam; PF00200; Disintegrin; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR ProDom; PD000664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR PROSITE; PS00215; ADAM\_MPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS00244; DISINTEGRIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.



DR PROSITE; PS01186; EGF\_2; FALSE NEG.  
KW EGF-like domain; Glycoprotein; Signal; Spermatogenesis; Transmembrane.  
FT SIGNAL 16 Potential.  
FT PROPEP 17 184 Potential.  
FT CHAIN 185 739 ADAM 18.  
FT DOMAIN 177 687 Extracellular (Potential).  
FT TRANSMEM 708 Potential.  
FT DOMAIN 709 739 Cytoplasmic (Potential).  
FT DOMAIN 177 389 Metalloprotease-like.  
FT DOMAIN 390 479 Disintegrin-like.  
FT DOMAIN 480 622 Cys-rich.  
FT DOMAIN 620 654 EGF-like.  
FT DISULFID 293 376 By similarity.  
FT DISULFID 335 360 By similarity.  
FT DISULFID 458 483 Potential.  
FT DISULFID 624 636 By similarity.  
FT DISULFID 630 642 By similarity.  
FT DISULFID 644 653 By similarity.  
FT CARBOHYD 36 36 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 76 76 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 122 122 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 149 149 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 156 156 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 177 177 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 294 294 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 359 359 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 465 465 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 561 561 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 611 611 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 625 625 N-linked (GlcNAc... ) (Potential).  
SQ SEQUENCE 739 AA; 82856 MW; 2D8BE9A975072CDD CRC64;

Query Match 39.7%; Score 1582.5; DB 1; Length 739;  
Best Local Similarity 41.9%; Pred. No. 3.6e-99;  
Matches 314; Conservative 128; Mismatches 275; Indels 33; Gaps 13;

Qy 3 VFLSLGLGRLMDSNFDLPVQITVPKIRSIIEKIEGIESQASKIVIEGKPYTNLMQK 62  
Db 4 LLALLTLGLRLQAHEGSGIPLHVTVPKIRSDNSEVSEKRWIYIITIDGQPYTLHLGKQ 63  
Qy 63 NPLPHNRVSYSGTGIMKPLDQFQNFCHVQGYIEGPKSVVMVSTCTGLRGVLPQENV 122  
Db 64 SFLPQNFVLYTYNETGSLHSVSPYFMHCHYQGYAAEFNSFVTLSCGLRGVLPQENI 123  
Qy 123 SYGIEPLESSVGFHVIYQVHKKADVSL--YNEKDIERSRLSPKLOSABEQPOFAK-- 177  
Db 124 SYGIEPLESSVGFHVIYQVHKKADVSL--YNEKDIERSRLSPKLOSABEQPOFAK-- 183  
Qy 178 -YIEMHVIKQLYNHMGSDTVVVAQKVFQIGLTNAIFVSNITIIILSSLELWIDENKI 236  
Db 184 QYLEIYIIVEKALVDYMGSEMMAVTQKIVQIGLVNTMFTQFKLTIVILSSLELWSNENQI 243  
Qy 237 ATTGEANELLHTELRLWKTSYLVLPBHDVAFLLVYREKSNVYCATFOGKMCNDANYAGVVL 296  
Db 244 STSGDADDILQRFIAWKRDYLLPBDIAYLLVYRKHPKYVGATFPCTVCNKSVDAGIAM 303  
Qy 297 HPRTISLESIAVLAQLLSLSMGITYDDINKCCQSGAVCMNPEAIHFSVGKIFSNCSFE 356  
Db 304 YPDAIGLEGFSVIAQLGLNVLGYTYDDITQCFCLRATCINMHEAVASGRKIFSNCSMH 363  
Qy 357 DFAHFISKQSLHNLQRLDPFPKQAVCGNAKLEAGEECDCGTEDCALIGETCCDIA 416  
Db 364 DYRYFVSKFTKCLQKLSNLQPLHONQFVPCNGGILESENEECDCGNKNEQCF--KKCCDYN 421  
Qy 417 TCRFKAGSNCAEGPCC-ENCLFMSEKERMCRSPF-ECOLPEYCNCGSASCENHYVQTGH 474  
Db 422 TCKLUGSVKSGSGPCCSKBELSIAGTFCRKSIDPECDFTCYCNGTSSNCVPDPTALNGR 481  
Qy 475 PCGLNQMICIDGVCMWGDQKCTDTDFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQC 534  
Db 482 LCKLGTAYCNGQCTTNDQCAKIFGKAQAPACFPCKEVLNHERENCGFNKSNQPLPC 541  
Qy 535 EADNLQCGKGLICKYVGVKFLQIPRATTIYANISGHLCTAV----EFASDHADSQMWIKD 590

Db 542 ERKDVLCGLKACVQPHKANKSDAQSTVYSYIQQHVCVSIATGSSMSRSDGTDN--AYVAD 599  
Qy 591 GTSCGSKNVCBNQRCVSSSYLVGYDC-TTDKCNDRGVCNKKKHCHCSASYLPDPSCVQSDL 649  
Db 600 GTMCGPEMYCVNKTCKRVHLMGYNCNATTKCKGKICNFGNCQCFPGHRRPDPCKFOFGS 659  
Qy 650 WPGSIDSGNFPFVPAIPARLPERRIENIYHSEKMRW---PFFLFIPFFIFCVLIAMV 706  
Db 660 -PGSIDDGNF-----QKSGDFYTEKGYNTHWNWFIILSCIFLFFIVFTTVI---P 708  
Qy 707 KVNFRQKKWRTE--DYSSDEQPESESEPKG 734  
Db 709 KRNEISKCNRENAEYNRNSVVSDDVG 738

RESULT 13  
Q8TC27  
ID Q8TC27 PRELIMINARY; PRT; 787 AA.  
AC Q8TC27;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE A disintegrin and metalloprotease domain 32 (Similar to MDC family).  
GN Name=ADAM32; ORFNames=UNQ5982;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=22389257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Maman A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX Strausberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX Strausberg R.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2287296; PubMed=12975309;  
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,



RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,  
 RA Vandlen R., Watanabe C., Wiedand D., Woods K., Xie M.H., Yansura D.,  
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P.;  
 RA "The secreted protein discovery initiative (SPDI), a large-scale  
 RT effort to identify novel human secreted and transmembrane proteins: a  
 RT bioinformatics assessment.";  
 RL Genome Res. 13:2265-2270(2003).  
 DR ENBL; BC026169; AAH26169.1; -  
 DR ENBL; BC028702; AAH28702.1; -  
 DR ENBL; BC030014; AAH30014.1; -  
 DR ENBL; BC030698; AAH30698.1; -  
 DR ENBL; BC034975; AAH34975.1; -  
 DR ENBL; AV358739; AA089099.1; -  
 DR HSPP; P30403; IN4Y.  
 DR MEROPS; M12.960; -.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001762; disintegrin.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR Pfam; PF00200; Disintegrin; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR ProDom; PD000664; Disintegrin; 1.  
 DR PROSITE; PS0215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS0214; DISINTEGRIN\_2; 1.  
 DR PROSITE; PS01186; EGF\_2; UNKNOWN\_1.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 DR PROSITE; EGF\_3; 1.  
 KW Integrin; Metalloprotease; Protease.  
 SQ SEQUENCE 787 AA; 87933 MW; 3D84CACFECC1A12E CRC64;

Query Match 38.3%; Score 1524.5; DB 2; Length 787;  
 Best Local Similarity 41.1%; Pred. No. 3.4e-95;  
 Matches 310; Conservative 134; Mismatches 259; Indels 51; Gaps 18;

QY 1 MWVLLSGLGLRMD-SNFDSPVQITVPEKIRSIIEKIEGIE---SQASKIVIEGKPYT 56  
 DB 4 LWL--LLAGLCGLLASRPGFQNSLLQIVPEIKIQTNTDSSEIEYEQISYIIPIDEKLYT 61  
 QY 57 VNLMOKNFLPHNFRVSYSGTGIMKPLDQDFONFCHYQVIEGPKSVVWVSTCTGLRGV 116  
 DB 62 VHLKQRYFLADNFMVILYN-QGSNNTYSSDIQTQCYQGNIEGYPDSMVLSTCGLRGI 120  
 QY 117 LQFNVSYGIEPLESSVGFHEVIYQVKHKADVSLYNEKDIESR---DLGFKLQASBPQ 173  
 DB 121 LQFNVSYGIEPLESAVEFQHVLYKLVKNEVDNIAIFIDRLSKQPMDDNFISEKSEPAV 180  
 QY 174 D--FAKVIEMHVIVKOLYNHMGSDTTVAQKVPQLIGLTNAIFVSNITILSSLELWI 231

DB 418 CDFRTRCVLKGAKCYKGLCKDCQILQSGVECRPKAHPEDIAENCNGSSPECPDITL 477  
 QY 471 QTGHPCLNQMWCIDGVCMVSGDKQCTDTTFOKEVEFGPSECYSHLNSKTYDVGNGIS-DS 529  
 DB 478 INGLSCNNKFKICYDGDCHDLDCESVFGKSGRNAPFACYEETQSOSDRFGNCRDRNN 537  
 QY 530 GYTQCEADNLQCGKLCKYVKGKLLQIPRATIIYANISGHLCTIAVEFASDHADSQMKWIK 589  
 DB 538 KYVFCGWRNLICGLRLVCTYPTKPFHQENGVDVIAFVRDVSVCITVQYKLRTPVDPDLAVK 597  
 QY 590 DGTSCSGNKNQRCNRQSVSSYL---GYDCTTDCNDRGVGNKKHCHCSASYLPPDCSVQ 646  
 DB 598 NGSQCIDGRVGNRECVESRIIKAHVC-SQSCSGHGVCDNRKCHCSFGYKPPNCOIR 656  
 QY 647 SLDWPGSGI-----DSGNFPVPAIPARLPERRYENIYHSPKMW--PFFLPIFFIIFCV 700  
 DB 657 S---KGFSIPPEEDMGSI-----MERASGKTENTWLLGLFLIALPILIV--- 696  
 QY 701 LIAIMVKNFORKWRTEDYSSDEQSESEPEPK 734  
 DB 697 -TTAIVLARKQLKKWFAKE---EEFPSSSEKSEG 726

RESULT 14  
 AAQ89099 PRELIMINARY; PRT; 787 AA.  
 AC AAQ89099;  
 DT 02-MAR-2004 (Tremblrel. 27, Created)  
 DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)  
 DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)  
 DE Similar to MDC family.  
 GN UNQ5982.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12975309;  
 RA Clark H.P., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,  
 RA Vandlen R., Watanabe C., Wiedand D., Woods K., Xie M.H., Yansura D.,  
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P.;  
 RA "The Secreted Protein Discovery Initiative (SPDI), a Large-Scale  
 RT Effort to Identify Novel Human Secreted and Transmembrane Proteins: A  
 RT Bioinformatics Assessment.";  
 RL Genome Res. 13:2265-2270(2003).  
 DR ENBL; AV358739; AA089099.1; -  
 SQ SEQUENCE 787 AA; 87933 MW; 3D84CACFECC1A12E CRC64;

Query Match 38.3%; Score 1524.5; DB 2; Length 787;  
 Best Local Similarity 41.1%; Pred. No. 3.4e-95;  
 Matches 310; Conservative 134; Mismatches 259; Indels 51; Gaps 18;

QY 1 MWVLLSGLGLRMD-SNFDSPVQITVPEKIRSIIEKIEGIE---SQASKIVIEGKPYT 56  
 DB 4 LWL--LLAGLCGLLASRPGFQNSLLQIVPEIKIQTNTDSSEIEYEQISYIIPIDEKLYT 61  
 QY 57 VNLMOKNFLPHNFRVSYSGTGIMKPLDQDFONFCHYQVIEGPKSVVWVSTCTGLRGV 116  
 DB 62 VHLKQRYFLADNFMVILYN-QGSNNTYSSDIQTQCYQGNIEGYPDSMVLSTCGLRGI 120  
 QY 117 LQFNVSYGIEPLESSVGFHEVIYQVKHKADVSLYNEKDIESR---DLGFKLQASBPQ 173  
 DB 121 LQFNVSYGIEPLESAVEFQHVLYKLVKNEVDNIAIFIDRLSKQPMDDNFISEKSEPAV 180  
 QY 174 D--FAKVIEMHVIVKOLYNHMGSDTTVAQKVPQLIGLTNAIFVSNITILSSLELWI 231

Db 181 PDLPFLYLEHIVVDKTLTYDYGSDSMIVTNKVIIEVLGANSMTQFKVTIVLSLELWS 240  
Qy 232 DENKIATTGEANELLHPLRWKTSYLVRPHDVAFLVYREKSNVVGATFGKWCNDANYA 291  
Db 241 DENKISTVGEADELLQKFLQWQSYLNRPHDIAIYLLIYMDYPRYLGAFFGTWCITRYS 300  
Qy 292 GGVVLHPTISLESIAVLAQLLSLGMITDINKKQCSGAVCIMMPEALHFGSVKIFS 351  
Db 301 AGVALYPKEITLFAFVITQMLALSIGSYDDPKKQCSSESTCIMMPEVQSVNGVKTFS 360  
Qy 352 NCSFEDFAHFISSKQSCQLHNPRLDPFFKQAOVCGNAKLRAEGECDCGTEQCALIGET 411  
Db 361 SCSLRSFQNFISNVGVKQLQKQWQK-KSPKPCVCGNRLGEGNEICDCGTEAQCG--PAS 417  
Qy 412 CCDIATCRFKAGSNCAEPCENCLFMSKRMCRP-SFEECDLPEYCNSSASCPENHYV 470  
Db 418 CCDFRICVLKDGAKYKGLCKCQILQSGVECPKHAHPCDIAENCGSSPECGPDITL 477  
Qy 471 QTGHPCGLNQCWICIDGVCMSGDQCTDTDFGKEVFGPSECYSHLNSKTDVSGNCGIS-DS 529  
Db 478 INGLSKNNKFCYDGDCHDLARCESVFGKSRNAPFACVEEIQSQSDRFGNCGDRNN 537  
Qy 530 GYTQCEADNLCQGLKICYVGFLLQIPRATIIYANISHGHCIAVEFASDHADSKWKIK 589  
Db 538 KYVFCGWRNLICGLVCTYPTKPFQHBNGDVIVAFVRDSCVITDVKLPRTVPDPLAVK 597  
Qy 590 DGTSCGSKVCRNORCVSSSVL---GYDCTTDKNDRGVGNKKNKCHCSASYLPDQSVQ 646  
Db 598 NGSQCDIGRVCNRECVESRIKASAVC-SQCSGSHGVCDNRNKKCHCSGPKPKPCOIR 656  
Qy 647 SLDLPGGSI----DSGNFPFVPAIPARLPERRYIENIYHSPMRV--PFFLIPFFIIFCV 700  
Db 657 S---KGFSIPEEDMGSI-----MERASGKTENTWLLGFLIALPILIV---696  
Qy 701 LIAIMVKNFQKRWRTEDVSSDQPSSEPKG 734  
Db 697 -TTAIVLARKQLKWKFAKE---EEFSPSEKSEG 726

RESULT 15  
Q8TC42 PRELIMINARY; PRT; 787 AA.  
AC Q8TC42;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE A disintegrin and metalloprotease domain 32.  
GN Name=ADAM32;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore S.I., Wang J., Hong F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Vallalon D.K., Muzny D.M., Soderstrom E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A.C., Young A.C., Shevchenko V., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.

"Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Strauberg R.,  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC026085; AAH26085.1; -  
DR HSPF; P30403; IN41.  
DR MEROPS; M12.960; -  
GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.  
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR006586; ADAM\_cysteine.  
DR InterPro; IPR001762; Disintegrin.  
DR InterPro; IPR000742; EGF 2.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR001590; Peptidase\_M12B.  
DR InterPro; IPR002870; Peptidase\_M12B\_N.  
DR Pfam; PF00200; Disintegrin; 1.  
DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
DR Pfam; PF01421; Repolysin; 1.  
DR ProDom; PD000664; Disintegrin; 1.  
DR SMART; SM00608; ACR; 1.  
DR SMART; SM00050; DISIN; 1.  
DR SMART; SM00381; EGF; 1.  
DR PROSITE; PS00215; ADAM\_MEPRO; 1.  
DR PROSITE; PS0214; DISINTEGRIN 2; 1.  
DR PROSITE; PS01186; EGF 2; UNKNOWN 1.  
DR PROSITE; PS50026; EGF 3; 1.  
KW Integrin; Metalloprotease; Protease.  
SQ SEQUENCE 787 AA; 87931 MW; CE0F54A50466B336 CRC64;  
Query Match 38.1%; Score 1519.5; DB 2; Length 787;  
Best Local Similarity 41.0%; Pred. No. 7.5e-95;  
Matches 309; Conservative 135; Mismatches 259; Indels 51; Gaps 18;  
Qy 1 MWVLLSGLGLRMD-SNFDLPLVQITVPEKTRSIKEGIE---SQASKIVIEGPKYT 56  
Db 4 LWL--LLAGLGLASRPQFQNSLLQIVPEKIQTNTDSSEIEYEIGIIVIDEKLYT 61  
Qy 57 VNLMOQNFLPHNFRVSYSGTGIMKPLDQDFQNFCHQYGYIEGPKSVVWVSTCTGLRGV 116  
Db 62 VHLKQRYFLADNFWIYLYN-QGSMYYSDDIQCYRGNIEGYPDMSVTLSTCSGLRGI 120  
Qy 117 LQFENVSYGIEPLESSVGFPHVYVQVKKKADVSLYNEKDIESR---DLSFKLQSAEPQ 173  
Db 121 LQFENVSYGIEPLESAVEFQHLVLYKLKVEDNDIAIFIDRGLKEQPMDDNIFISEKSPAV 180  
Qy 174 D--FAKIEHVIYEVKQYLNHMGSDTIVVAQKVPQLIGLTNAIFVSNITIIISLLEWI 231  
Db 181 PDLFPLYLEHIVVDKTLTYDYGSDSMIVTNKVIIEVLGANSMTQFKVTIVLSLELWS 240  
Qy 232 DENKIATTGEANELLHPTISLESIAVLAQLLSLGMITDINKKQCSGAVCIMMPEALHFGSVKIFS 291  
Db 241 DENKISTVGEADELLQKFLQWQSYLNRPHDIAIYLLIYMDYPRYLGAFFGTWCITRYS 300  
Qy 292 GGVVLHPTISLESIAVLAQLLSLGMITDINKKQCSGAVCIMMPEALHFGSVKIFS 351  
Db 301 AGVALYPKEITLFAFVITQMLALSIGSYDDPKKQCSSESTCIMMPEVQSVNGVKTFS 360  
Qy 352 NCSFEDFAHFISSKQSCQLHNPRLDPFFKQAOVCGNAKLRAEGECDCGTEQCALIGET 411  
Db 361 SCSLRSFQNFISNVGVKQLQKQWQK-KSPKPCVCGNRLGEGNEICDCGTEAQCG--PAS 417  
Qy 412 CCDIATCRFKAGSNCAEPCENCLFMSKRMCRP-SFEECDLPEYCNSSASCPENHYV 470  
Db 418 CCDFRICVLKDGAKYKGLCKCQILQSGVECPKHAHPCDIAENCGSSPECGPDITL 477  
Qy 471 QTGHPCGLNQCWICIDGVCMSGDQCTDTDFGKEVFGPSECYSHLNSKTDVSGNCGIS-DS 529

```
Db 478 INGLSCNKNKFCYDGDCHDLDAECESVFGKGSNAPFACYEEIQSQSDRFGNCGRDRNN 537
Qy 530 GYTQCEADNLQCGKLIQKLYGKFLQIPRATIIYANISGHLCIAVEFASDHADSQKMWIK 589
Db 538 KYVFCGWRNLICGRLVCTYPTKRFHQENGDIYAFVRDSCITVDYKLPRTVPDPLAVK 597
Qy 590 DGTSCGSNKYCRNORCVSSYL---GYDCTTDKNDRGVCNNKXKHCHCSASYLPPDCSVQ 646
Db 598 NGSQCDIGRVCNRECVESRIIKASAHVC-SQCGSGHGVCDNRNKHCHCSFGYKPPNCOIR 656
Qy 647 SDLWPGGSI---DSGNFPPVAIPARLPERRYENIYHVKPMRW--PFFLFIPFFIIFCV 700
Db 657 S---KGFSIFPEEDMGSI-----MERASGKTENTWLLGFLIALPLIV--- 696
Qy 701 LIAMVKVNFQRKWRTEYSSDEQPESEPEPKG 734
Db 697 -TTAIVLARQLKWKFAKE---BEFPPSESKSEG 726
```

Search completed: January 10, 2005, 22:27:44  
Job time : 204 secs

*This Page Blank (uspto)*



```
Db 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFE 120
QY 121 NVSYGTEPLESSVGFHVIVQVHKHKAADVLSYNEKDIESRDLSPKLSAEPQODFAKYIE 180
Db 121 NVSYGTEPLESSVGFHVIVQVHKHKAADVLSYNEKDIESRDLSPKLSAEPQODFAKYIE 180
QY 181 MHVIVEKQLYNHMGSDTTVVAQKVFQLIGLTNAIFVSNITIISSLELWIDENKIATTG 240
Db 181 MHVIVEKQLYNHMGSDTTVVAQKVFQLIGLTNAIFVSNITIISSLELWIDENKIATTG 240
QY 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMCDANVAGGVLLHPT 300
Db 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMCDANVAGGVLLHPT 300
QY 301 ISLESIAVILAOQLLSLMSGITYDDINKCQCSGAVCIMNPEAIHFSGVKIFSNCSFEDFAH 360
Db 301 ISLESIAVILAOQLLSLMSGITYDDINKCQCSGAVCIMNPEAIHFSGVKIFSNCSFEDFAH 360
QY 361 FISKQKSQCLHNPRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420
Db 361 FISKQKSQCLHNPRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420
QY 421 KAGSNCAEGPCENCLFMSKERMCRPSFECDLPEYCNSSASCPENHYVQTGHPCGLNQ 480
Db 421 KAGSNCAEGPCENCLFMSKERMCRPSFECDLPEYCNSSASCPENHYVQTGHPCGLNQ 480
QY 481 WICIDGVCMSGDKQCTDTFGKEVFGPSECYSHLNSKTDVSGNCGISDSGVTQCEADNLQ 540
Db 481 WICIDGVCMSGDKQCTDTFGKEVFGPSECYSHLNSKTDVSGNCGISDSGVTQCEADNLQ 540
QY 541 CGKLI CKYVGFLLQIPRATIIYANISGHLICIAVEFASDHADSQKWKIDGTSCGSKVC 600
Db 541 CGKLI CKYVGFLLQIPRATIIYANISGHLICIAVEFASDHADSQKWKIDGTSCGSKVC 600
QY 601 RNQRCVSSSYLVGDDCTTDKCNDRGVNCKKHCHCSASYLPPDCSVQSDLPFGSIDSNGF 660
Db 601 RNQRCVSSSYLVGDDCTTDKCNDRGVNCKKHCHCSASYLPPDCSVQSDLPFGSIDSNGF 660
QY 661 PPVAIPARLPERRYENIYHSKPMRWPFLLFIPFFIIFCVLIIAMVKNFORKKWRTEDY 720
Db 661 PPVAIPARLPERRYENIYHSKPMRWPFLLFIPFFIIFCVLIIAMVKNFORKKWRTEDY 720
QY 721 SSDEQPESESEPKG 734
Db 721 SSDEQPESESEPKG 734
```

## RESULT 2

```
US-10-205-823-10
; Sequence 10, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsley, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR FILING DATE: 2002-07-25
; PRIOR FILING DATE: 2001-07-25
; PRIOR FILING DATE: 2001-07-25
; PRIOR FILING DATE: 2001-08-22
```

```
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-205-823-10
```

```
Query Match 100.0%; Score 3984; DB 14; Length 734;
Best Local Similarity 100.0%; Pred. No. 3.5e-311;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWVLFLLSGLGGLRMDSNFDSLPVQITVPEKIRSIIEKIGESQASYKIVIEGKPYTVNLM 60
Db 1 MWVLFLLSGLGGLRMDSNFDSLPVQITVPEKIRSIIEKIGESQASYKIVIEGKPYTVNLM 60
QY 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFE 120
Db 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFE 120
QY 121 NVSYGTEPLESSVGFHVIVQVHKHKAADVLSYNEKDIESRDLSPKLSAEPQODFAKYIE 180
Db 121 NVSYGTEPLESSVGFHVIVQVHKHKAADVLSYNEKDIESRDLSPKLSAEPQODFAKYIE 180
QY 181 MHVIVEKQLYNHMGSDTTVVAQKVFQLIGLTNAIFVSNITIISSLELWIDENKIATTG 240
Db 181 MHVIVEKQLYNHMGSDTTVVAQKVFQLIGLTNAIFVSNITIISSLELWIDENKIATTG 240
QY 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMCDANVAGGVLLHPT 300
Db 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMCDANVAGGVLLHPT 300
QY 301 ISLESIAVILAOQLLSLMSGITYDDINKCQCSGAVCIMNPEAIHFSGVKIFSNCSFEDFAH 360
Db 301 ISLESIAVILAOQLLSLMSGITYDDINKCQCSGAVCIMNPEAIHFSGVKIFSNCSFEDFAH 360
QY 361 FISKQKSQCLHNPRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420
Db 361 FISKQKSQCLHNPRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420
QY 421 KAGSNCAEGPCENCLFMSKERMCRPSFECDLPEYCNSSASCPENHYVQTGHPCGLNQ 480
Db 421 KAGSNCAEGPCENCLFMSKERMCRPSFECDLPEYCNSSASCPENHYVQTGHPCGLNQ 480
QY 481 WICIDGVCMSGDKQCTDTFGKEVFGPSECYSHLNSKTDVSGNCGISDSGVTQCEADNLQ 540
Db 481 WICIDGVCMSGDKQCTDTFGKEVFGPSECYSHLNSKTDVSGNCGISDSGVTQCEADNLQ 540
QY 541 CGKLI CKYVGFLLQIPRATIIYANISGHLICIAVEFASDHADSQKWKIDGTSCGSKVC 600
Db 541 CGKLI CKYVGFLLQIPRATIIYANISGHLICIAVEFASDHADSQKWKIDGTSCGSKVC 600
QY 601 RNQRCVSSSYLVGDDCTTDKCNDRGVNCKKHCHCSASYLPPDCSVQSDLPFGSIDSNGF 660
Db 601 RNQRCVSSSYLVGDDCTTDKCNDRGVNCKKHCHCSASYLPPDCSVQSDLPFGSIDSNGF 660
QY 661 PPVAIPARLPERRYENIYHSKPMRWPFLLFIPFFIIFCVLIIAMVKNFORKKWRTEDY 720
Db 661 PPVAIPARLPERRYENIYHSKPMRWPFLLFIPFFIIFCVLIIAMVKNFORKKWRTEDY 720
QY 721 SSDEQPESESEPKG 734
Db 721 SSDEQPESESEPKG 734
```

## RESULT 3

US-10-473-603-19

```
; Sequence 19, Application US/10473603
; Publication No. US20040235066A1
; GENERAL INFORMATION:
; APPLICANT: OLD, Lloyd
; APPLICANT: SCANLAN, Matthew
; APPLICANT: CHEN, Yao-Tseng
; TITLE OF INVENTION: CANCER-TESTIS ANTIGENS
; FILE REFERENCE: L0461.701550S00
; CURRENT FILING DATE: 2003-09-30
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: PCT/US02/09808
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 10/054,683
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/280,718
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/285,154
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/327,432
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 734
; TYPE: PR1
; ORGANISM: Homo Sapiens
; US-10-473-603-19

Query Match      100.0%; Score 3984; DB 17; Length 734;
Best Local Similarity 100.0%; Pred. No. 3.5e-311;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWVLLSGLGRLMDSNFDSPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNL 60
DB 1 MWVLLSGLGRLMDSNFDSPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNL 60
QY 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQYIEGPKSVVMVSTCTGLRGVLQPE 120
DB 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQYIEGPKSVVMVSTCTGLRGVLQPE 120
QY 121 NVSYGIEPLESSVGFHEVIYQVKKKADVSLYNEKDIESRDLSPKLOSAPQODFAKYE 180
DB 121 NVSYGIEPLESSVGFHEVIYQVKKKADVSLYNEKDIESRDLSPKLOSAPQODFAKYE 180
QY 181 MHVIVEKQLYNHMGSDTTVVAQVQLIGLTNAIFVSFNITIISSLELWIDENKIATTG 240
DB 181 MHVIVEKQLYNHMGSDTTVVAQVQLIGLTNAIFVSFNITIISSLELWIDENKIATTG 240
QY 241 EANELHTFLRWKTSYLVRPHDVAFLVYREKSNVVGATFQGMKCDANYAGGVVLHPRT 300
DB 241 EANELHTFLRWKTSYLVRPHDVAFLVYREKSNVVGATFQGMKCDANYAGGVVLHPRT 300
QY 301 ISLESIAVILAOLLSLMSGITYDDINKCQCSGAVCIIMNPEAIIHFSGVKIFSNCSFEDFAH 360
DB 301 ISLESIAVILAOLLSLMSGITYDDINKCQCSGAVCIIMNPEAIIHFSGVKIFSNCSFEDFAH 360
QY 361 FISKQKSOCLHNPRLDPPFPKQAVCGNAKLEAGECDCTEQDCALIGTCCDIATCRF 420
DB 361 FISKQKSOCLHNPRLDPPFPKQAVCGNAKLEAGECDCTEQDCALIGTCCDIATCRF 420
QY 421 KAGSNCAEGPCCNCLFMSKERMCRPSFEBCDLPYECNGSSASCENHYVQTGHPGCLNQ 480
DB 421 KAGSNCAEGPCCNCLFMSKERMCRPSFEBCDLPYECNGSSASCENHYVQTGHPGCLNQ 480
QY 481 WICIDGVCMSGDKOCTDTTFCKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTCEADNLQ 540
DB 481 WICIDGVCMSGDKOCTDTTFCKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTCEADNLQ 540
QY 541 CGKLI CKYGVKFLQLPRATIIIVANTISGHLICIAVEFASDHADSQKMWIKDGTSCGNKVC 600
DB 541 CGKLI CKYGVKFLQLPRATIIIVANTISGHLICIAVEFASDHADSQKMWIKDGTSCGNKVC 600
QY 601 RNQRCSVSSYGLDYCTTDCNDRGVNCKKCHCSASYLPDPCSVQSDLWPGGSIDSGNF 660
DB 601 RNQRCSVSSYGLDYCTTDCNDRGVNCKKCHCSASYLPDPCSVQSDLWPGGSIDSGNF 660
QY 661 PPVAIPARLPERRYIENIYHSKPMRPPFFLFIFFFIICVLIAMVKNVFNQKWKRTEDY 720
DB 661 PPVAIPARLPERRYIENIYHSKPMRPPFFLFIFFFIICVLIAMVKNVFNQKWKRTEDY 720
QY 721 SSDEQPESESEPKG 734
DB 721 SSDEQPESESEPKG 734

RESULT 4
US-10-205-823-8
; Sequence 8, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-10-205-823-8

Query Match      99.6%; Score 3969.5; DB 14; Length 735;
Best Local Similarity 99.7%; Pred. No. 5.1e-310;
Matches 733; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MW-VLLSGLGRLMDSNFDSPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNL 59
DB 1 MWVLLSGLGRLMDSNFDSPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNL 60
QY 60 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQYIEGPKSVVMVSTCTGLRGVLQF 119
DB 61 MQKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQYIEGPKSVVMVSTCTGLRGVLQF 120
QY 120 ENVSYGIEPLESSVGFHEVIYQVKKKADVSLYNEKDIESRDLSPKLOSAPQODFAKYE 179
DB 121 ENVSYGIEPLESSVGFHEVIYQVKKKADVSLYNEKDIESRDLSPKLOSAPQODFAKYE 180
QY 180 EMHVI VEKQLYNHMGSDTTVVAQVQLIGLTNAIFVSFNITIISSLELWIDENKIATT 239
DB 181 EMHVI VEKQLYNHMGSDTTVVAQVQLIGLTNAIFVSFNITIISSLELWIDENKIATT 240
QY 240 GEANELHTFLRWKTSYLVRPHDVAFLVYREKSNVVGATFQGMKCDANYAGGVVLHPR 299
DB 240 GEANELHTFLRWKTSYLVRPHDVAFLVYREKSNVVGATFQGMKCDANYAGGVVLHPR 299
```

```

Db 241 GEANELHFLRWKTSYLVLRPHDVAFLLYREKSNVVGATFGKMCNDANYAGGVULHPR 300
QY 300 TISLESIAVLAQLLSMGITVDDINKKCCSGAVCIMNPEAIHFGVKIFSNCSFEDFA 359
Db 301 TISLESIAVLAQLLSMGITVDDINKKCCSGAVCIMNPEAIHFGVKIFSNCSFEDFA 360
QY 360 HFTSKOKSOCLHNPRLDPFFKQAVCGNAKLEAGEECDCGTODCALIGETCCDIATCR 419
Db 361 HFTSKOKSOCLHNPRLDPFFKQAVCGNAKLEAGEECDCGTODCALIGETCCDIATCR 420
QY 420 FKAGSNCAEGPCCECENCLFMSKERMCRPSFECDLPEYCNCGSSASCSPENHYVQTGHPCGLN 479
Db 421 FKAGSNCAEGPCCECENCLFMSKERMCRPSFECDLPEYCNCGSSASCSPENHYVQTGHPCGLN 480
QY 480 QWICIDGVCMSGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNL 539
Db 481 QWICIDGVCMSGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNL 540
QY 540 QCGKLICKYVKGFLQIPRATIIYANISGHLCAVEFASDHADSQKWIKDGTSCGSNKV 599
Db 541 QCGKLICKYVKGFLQIPRATIIYANISGHLCAVEFASDHADSQKWIKDGTSCGSNKV 600
QY 600 CRNQRVCSSYLGVDCTTDCNDRGVCCNKKHCHCSASYLPDPCSVOQSDLWPGGSDSGN 659
Db 601 CRNQRVCSSYLGVDCTTDCNDRGVCCNKKHCHCSASYLPDPCSVOQSDLWPGGSDSGN 660
QY 660 FPPVAIPARIPERRIENIYHSPMRWPFLLFPFFIIFCVLIAIMVKVNFQKRWRTED 719
Db 661 FPPVAIPARIPERRIENIYHSPMRWPFLLFPFFIIFCVLIAIMVKVNFQKRWRTED 720
QY 720 YSSDEQSESEPKG 734
Db 721 YSSDEQSESEPKG 735

RESULT 5
US-10-205-823-6
; Sequence 6, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsley, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 753
; TYPE: PRT

; ORGANISM: Homo sapiens
US-10-205-823-6
Query Match 97.8%; Score 3898; DB 14; Length 753;
Best Local Similarity 99.0%; Pred. No. 2.9e-304; Indels 0; Gaps 0;
Matches 718; Conservative 1; Mismatches 6;

QY 10 LGGLRMSDNFSLPQVITVPEKIRSIIRIGIEISQASYKIVIEGKPYTNLMQKNFLPHNF 69
Db 29 LGKLYTDQDFDSLPAQITVPEKIRSIIRIGIEISQASYKIVIEGKPYTNLMQKNFLPHNF 88
QY 70 RYVSYSGTGIMKPLDQDFONFCHYQYIBGYPKSVVMSTCTGLRGVLOFENVSGIEPL 129
Db 89 RYVSYSGTGIMKPLDQDFONFCHYQYIBGYPKSVVMSTCTGLRGVLOFENVSGIEPL 148
QY 130 ESSVGFPEHVIYQVKKKADVSLYNEKDIERSRLSFKLQSAEPQODPAKVIEMHVI VEKQL 189
Db 149 ESSVGFPEHVIYQVKKKADVSLYNEKDIERSRLSFKLQSAEPQODPAKVIEMHVI VEKQL 208
QY 190 YNHMGSDTTTVAQKVFQLLGLTNALFVSNITIIILSSLELWIDENKIATGSEANELLHTF 249
Db 209 YNHMGSDTTTVAQKVFQLLGLTNALFVSNITIIILSSLELWIDENKIATGSEANELLHTF 268
QY 250 LRWKTSYLVLRPHDVAFLLYREKSNYVCATFGKMCNDANYAGGVULHPRTTISLESIAVI 309
Db 269 LRWKTSYLVLRPHDVAFLLYREKSNYVCATFGKMCNDANYAGGVULHPRTTISLESIAVI 328
QY 310 LAQLLSLSMGITVDDINKKCCSGAVCIMNPEAIHFGSVKIFSNCSFEDFAHFSKQSQ 369
Db 329 LAQLLSLSMGITVDDINKKCCSGAVCIMNPEAIHFGSVKIFSNCSFEDFAHFSKQSQ 388
QY 370 LHNQPRLDPPFKQAVCGNAKLEAGEECDCGTQDCCALIGETCCDIATCRFKAGSNCAEG 429
Db 389 LHNQPRLDPPFKQAVCGNAKLEAGEECDCGTQDCCALIGETCCDIATCRFKAGSNCAEG 448
QY 430 PCCECLFMSKERMCRPSFECDLPEYCNCGSSASCSPENHYVQTGHPCGLNQCICIDGVC 489
Db 449 PCCECLFMSKERMCRPSFECDLPEYCNCGSSASCSPENHYVQTGHPCGLNQCICIDGVC 508
QY 490 SGDQKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCKLICKYV 549
Db 509 SGDQKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCKLICKYV 568
QY 550 GKFLQIPRATIIYANISGHLCAVEFASDHADSQKWIKDGTSCGSNKVCRNQRVCSS 609
Db 569 GKFLQIPRATIIYANISGHLCAVEFASDHADSQKWIKDGTSCGSNKVCRNQRVCSS 628
QY 610 YLGYDCTTDCNDRGVCCNKKHCHCSASYLPDPCSVOQSDLWPGGSDSGNPPVAIPARL 669
Db 629 YLGYDCTTDCNDRGVCCNKKHCHCSASYLPDPCSVOQSDLWPGGSDSGNPPVAIPARL 688
QY 670 PERRYIENIYHSPMRWPFLLFPFFIIFCVLIAIMVKVNFQKRWRTEDYSSDEQSEPE 729
Db 689 PERRYIENIYHSPMRWPFLLFPFFIIFCVLIAIMVKVNFQKRWRTEDYSSDEQSEPE 748
QY 730 SEPKG 734
Db 749 SEPKG 753

RESULT 6
US-10-227-884-90
; Sequence 90, Application US/10227884
; Publication No. US20030027988A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Deenoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
```



APPLICANT: Stephan, Jean-Philippe F.  
APPLICANT: Watanabe, Colin L.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3530P1C79  
CURRENT APPLICATION NUMBER: US/10/227,884  
CURRENT FILING DATE: 2002-08-26  
PRIOR APPLICATION NUMBER: 10/119,480  
PRIOR FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063549  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/064103  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/069873  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086392  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089905  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090691  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/095302  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/095318  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/095916  
PRIOR FILING DATE: 1998-08-10  
PRIOR APPLICATION NUMBER: 60/096146  
PRIOR FILING DATE: 1998-08-11  
PRIOR APPLICATION NUMBER: 60/096791  
PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 60/097986  
PRIOR FILING DATE: 1998-08-26  
PRIOR APPLICATION NUMBER: 60/098544  
PRIOR FILING DATE: 1998-08-31  
PRIOR APPLICATION NUMBER: 60/099596  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 60/099598  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 60/099803  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099811  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099812  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099816  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/100038  
PRIOR FILING DATE: 1998-09-11  
PRIOR APPLICATION NUMBER: 60/100385  
PRIOR FILING DATE: 1998-09-15  
PRIOR APPLICATION NUMBER: 60/100390  
PRIOR FILING DATE: 1998-09-15  
PRIOR APPLICATION NUMBER: 60/100627  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: 60/100848  
PRIOR FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: 60/100919  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/101477  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101738  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101741  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101786  
PRIOR FILING DATE: 1998-09-25  
PRIOR APPLICATION NUMBER: 60/101916  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101922  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/106178  
PRIOR FILING DATE: 1998-10-28  
PRIOR APPLICATION NUMBER: 60/106248  
PRIOR FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: 60/106464  
PRIOR FILING DATE: 1998-10-30  
PRIOR APPLICATION NUMBER: 60/106905  
PRIOR FILING DATE: 1998-11-03  
PRIOR APPLICATION NUMBER: 60/108787  
PRIOR FILING DATE: 1998-11-17  
PRIOR APPLICATION NUMBER: 60/108801  
PRIOR FILING DATE: 1998-11-17  
PRIOR APPLICATION NUMBER: 60/108849  
PRIOR FILING DATE: 1998-11-18  
PRIOR APPLICATION NUMBER: 60/112422  
PRIOR FILING DATE: 1998-12-15  
PRIOR APPLICATION NUMBER: 60/113296  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/113605  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/113621  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/115558  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115565  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115733  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/119549  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 60/123618  
PRIOR FILING DATE: 1999-03-10  
PRIOR APPLICATION NUMBER: 60/125259  
PRIOR FILING DATE: 1999-03-19  
PRIOR APPLICATION NUMBER: 60/125775  
PRIOR FILING DATE: 1999-03-23  
PRIOR APPLICATION NUMBER: 60/126773  
PRIOR FILING DATE: 1999-03-29

```

; PRIOR APPLICATION NUMBER: 60/127887
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

Query Match      38.3%; Score 1524.5; DB 14; Length 787;
Best Local Similarity 41.1%; Pred. No. 2.1e-113;
Matches 310; Conservative 134; Mismatches 259; Indels 51; Gaps 18;

QY 1 MWVFLSLGGLRMD-SNFDLSLVQIVTPEKIRSIIEKIE-----SQASYKIVIEGPKYT 56
DB 4 LWL--LLAGLCGLASRPFQNSLLQIVPEKIQTNTNDSSEIEYEQISYIPIDEKLYT 61
QY 57 VNLWKNFLPHNRYVSYSGTGIMKPLDQFQNECHYGVTEGYPKSVVMVSTCTGLRGV 116
DB 62 VHLQRYFLADNFMILYN-QGSNWTSSDIQTQCYGNGIEGYPDSMTLSTCSGURGI 120
QY 117 LQENVSYGIPELSSVGFHVIYQVKKHADVSLYNEKIDIESR---DLSPKQLQASBPQ 173
DB 121 LQENVSYGIPELSEAVFQHVLYKLNEDNDIAIFIDRLSKQPMDDNFISEKSPAV 180
QY 174 D--FAKVIEMHVIKOLYNHMGSDTTTVAQKQFQLGLNAINPVSNITILSSLELWI 231
DB 181 PDLFPLYLEHIVVDKTLDYWGSDSMTVNTKVTIEIVGLANSMTQPKVTIVILSSLEWS 240
QY 232 DENKIATGEGANELLHTFLWKTSYLVLRHDVAFLLVYREKSNVYCATFGKMCNDANYA 291
DB 241 DENKISTVGADELQKFLWKQSYLNLRLPHDIAYLITIMDYPRYLGAFFGTWCITRYS 300
QY 292 GGVLHPRTISLESIAVILAQLLSLMSGITVDDINKQCCSGAVCIMNPEAIHFSGVKIFS 351
DB 301 AGVALYKPEITLEAFAVITQMLALSIGSYDDPKKQCCSESTCIMNPEVQSNVKTFS 360
QY 352 NCSFEDPAHFISKQSKOCLNHPRLDPFFKQQAQVCGNAKUEAGEECDCGTEQDICALIGET 411

; 361 SCSLRSFQNFISNVGVKCLQNKPMQK-KSPKPVCGNRLGREGNEICDCGTEAQCG--PAS 417
; 412 CCDIATCRFKAGSNCAEGPCCNCLFMSKERMCRP-SFEBECDLPEYCNCGSSASCENHYV 470
; 418 CCDFRITCVLKGAKCYKGLCKCKCQILQSGVECRPKAHPCEDIAENCNGSSPEGGPITL 477
; 471 QTGHPGGLNQCICIDGVCMGSGDKQCTDTFQKEVEFGPSECYSHLNSKTDVSGNCGIS-DS 529
; 478 INGLSKNNKFCYDGDGDCHDLDAECESVFGKSGRNAPFAFYEEIQSQSDRFGNCGRDRNN 537
; 530 GYTQCADNLCQKLI CKYVKGKELLQIPRATIIYANISGHLCTIAVEFASHADSQKWKIK 589
; 538 KYVFCGWRNLICGRLVCTYPTKRPFFHQENGVDVYAFVRDSVCITVDYKLPRTVDPDLAVK 597
; 590 DGTSCGSKNVCNRQCVSSSYL---GYDCTTDKNDRGVCNNKHCHCSASLYLPDCSVQ 646
; 598 NGSCQDIGRVNRECVESRIIKASHVC-SQCSGHGVCDNRKCHCSGYAPPNCQIR 656
; 647 SLDLWPGGSI-----DSGNFFPVAIPARLPERRYIENIYHSKPMRW--PFFLFIPIFFICV 700
; 657 S---KGFSIFPEEDMGSI-----MERASGKTENTWLLGLFLIALPILIV--- 696
; 701 LIAMKVNFORKKNWRTEDYSSDEQSESESPKG 734
; 697 -TTAIVLARKQLKXFAKE---EEPPSSSEKSEG 726

RESULT 7
US-10-230-163-90
; Sequence 90, Application US/10230163
; Publication No. US2003003635A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Deenoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C96
; CURRENT APPLICATION NUMBER: US/10/230,163
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
```

;	PRIOR FILING DATE:	1998-09-24	
;	PRIOR APPLICATION NUMBER:	60/101922	
;	PRIOR FILING DATE:	1998-09-24	
;	PRIOR APPLICATION NUMBER:	60/106178	
;	PRIOR FILING DATE:	1998-10-28	
;	PRIOR APPLICATION NUMBER:	60/106248	
;	PRIOR FILING DATE:	1998-10-29	
;	PRIOR APPLICATION NUMBER:	60/106464	
;	PRIOR FILING DATE:	1998-10-30	
;	PRIOR APPLICATION NUMBER:	60/106905	
;	PRIOR FILING DATE:	1998-11-03	
;	PRIOR APPLICATION NUMBER:	60/108787	
;	PRIOR FILING DATE:	1998-11-17	
;	PRIOR APPLICATION NUMBER:	60/108801	
;	PRIOR FILING DATE:	1998-11-17	
;	PRIOR APPLICATION NUMBER:	60/108849	
;	PRIOR FILING DATE:	1998-11-18	
;	PRIOR APPLICATION NUMBER:	60/112422	
;	PRIOR FILING DATE:	1998-12-15	
;	PRIOR APPLICATION NUMBER:	60/113296	
;	PRIOR FILING DATE:	1998-12-22	
;	PRIOR APPLICATION NUMBER:	60/113605	
;	PRIOR FILING DATE:	1998-12-23	
;	PRIOR APPLICATION NUMBER:	60/113621	
;	PRIOR FILING DATE:	1998-12-23	
;	PRIOR APPLICATION NUMBER:	60/115558	
;	PRIOR FILING DATE:	1999-01-12	
;	PRIOR APPLICATION NUMBER:	60/115565	
;	PRIOR FILING DATE:	1999-01-12	
;	PRIOR APPLICATION NUMBER:	60/115733	
;	PRIOR FILING DATE:	1999-01-12	
;	PRIOR APPLICATION NUMBER:	60/119549	
;	PRIOR FILING DATE:	1999-02-10	
;	PRIOR APPLICATION NUMBER:	60/123618	
;	PRIOR FILING DATE:	1999-03-10	
;	PRIOR APPLICATION NUMBER:	60/125259	
;	PRIOR FILING DATE:	1999-03-19	
;	PRIOR APPLICATION NUMBER:	60/125775	
;	PRIOR FILING DATE:	1999-03-23	
;	PRIOR APPLICATION NUMBER:	60/126773	
;	PRIOR FILING DATE:	1999-03-29	
;	PRIOR APPLICATION NUMBER:	60/127887	
;	PRIOR FILING DATE:	1999-04-05	
;	PRIOR APPLICATION NUMBER:	60/130232	
;	PRIOR FILING DATE:	1999-04-21	
;	PRIOR APPLICATION NUMBER:	60/131022	
;	PRIOR FILING DATE:	1999-04-26	
;	PRIOR APPLICATION NUMBER:	60/131270	
;	PRIOR FILING DATE:	1999-04-27	
;	PRIOR APPLICATION NUMBER:	60/131291	
;	PRIOR FILING DATE:	1999-04-27	
;	PRIOR APPLICATION NUMBER:	60/131445	
;	PRIOR FILING DATE:	1999-04-28	
;	PRIOR APPLICATION NUMBER:	60/134287	
;	PRIOR FILING DATE:	1999-05-14	
;	PRIOR APPLICATION NUMBER:	60/140650	
;	PRIOR FILING DATE:	1999-06-22	
;	PRIOR APPLICATION NUMBER:	60/140723	
;	PRIOR FILING DATE:	1999-06-22	
;	PRIOR APPLICATION NUMBER:	60/141037	
;	PRIOR FILING DATE:	1999-06-23	
;	PRIOR APPLICATION NUMBER:	60/144758	
;	PRIOR FILING DATE:	1999-07-20	
;	PRIOR APPLICATION NUMBER:	60/145698	
;	PRIOR FILING DATE:	1999-07-26	
;	PRIOR APPLICATION NUMBER:	60/146222	
;	PRIOR FILING DATE:	1999-07-28	
;	PRIOR APPLICATION NUMBER:	60/146963	
;	PRIOR FILING DATE:	1999-08-03	
;	PRIOR APPLICATION NUMBER:	60/149320	
;	PRIOR FILING DATE:	1999-08-17	
;	PRIOR APPLICATION NUMBER:	60/149638	
;	PRIOR FILING DATE:	1999-08-17	

```
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

Query Match      38.3%; Score 1524.5; DB 14; Length 787;
Best Local Similarity 41.1%; Pred. No. 2.1e-113;
Matches 310; Conservative 134; Mismatches 259; Indels 51; Gaps 18;

Qy 1 MWVLFLLSGGLRMD-SNFDLSPVQITVPEKIRSIIEKIEGIE---SQASYKIVIEGKPYT 56
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 LWL--LLAGLCGLASRPGFQNSLLQIVPEKIQTNDSEIEYEQISYIIPIDEKLYT 61

Qy 57 VNLMOKNFLPHNPRVYSYSGTGIMKPLDDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGV 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 VHLKQRYFLADNFMILYN-QGSMNTYSSDIQTQCYQGNIEGYPDSMVTLCSTGSLRGI 120

Qy 117 LQFENVSYGIEPLESSVGFHEHYQVKKHKKADVSLYNEKDIESR---DLSPKLOQSAEPQ 173
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 LQFENVSYGIEPLESAVEFQHVLYLKLKNEQNDIAIFIDRSLSKEQPMDDNIFISEKSEPAV 180

Qy 174 D--FAKYEIMHVIVEKQLYNHMGSDTTVAQKVFQILGTNAIFVSNITIISSLELWI 231
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 PDLFPLYLEHMHVVDKTYDYGSDSMIVTNKVIEIVGLANSMTQPKVTIVLSLELWS 240

Qy 232 DENKIATTEGANELLHTFLRWKTSYLVLRPHDVAFLVREKSNYVGTGQKMCADANYA 291
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 DENKISTVGEADELLQKFLWKQSYLNLPRHDIAVLLIYMDYPRYLGAVPFGTMCITRYS 300

Qy 292 GGVLHPRTISLESIAVILAQLLSMGITYDDINKQCCSGAVCIMNPEAHFSGVKIFS 351
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 AGVALYPKEITLEAFVAVITQMLALSIGISYDDPKKQCSESTCIMNPEVQVNGVKTFS 360

Qy 352 NCSEFEDPAHISKOKSLQNLQPLDFFQOQAVCGNAKLEAGBECDCGTEDQCALIGET 411
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361 SCSLRSPQNFISNVGVKCLQNKPMQK-KSPKPYCGNGRLEGNICDCGTEAQCG--PAS 417

Qy 412 CCDIATCRFKAGSNCAEGPCENCLFWSKERMCRP-SFECDLPEYKNGSSASCENHYV 470
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
418 CCDFRIVLDGAKYKGLCKQCOILQSGVECPKPAHPECDIAENCGSSPCEGPDITL 477

Qy 471 QTHGPCGLNOMICIDGCMGSDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGIS-DS 529
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
478 INGLSCKNNKFCYDGDCHLDARCESVFGKGSRNAPFACVEETQSQSDRFGNCGDRNN 537

Qy 530 GYTQCEADNLCQKLIKCYGKFLQIIPRATIIYANISGHLCTIAVEFASDHASQKWIK 589
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
538 KYVFCGWRNLICGLRLVCTYPRKPFHOENGDDVIYAFVDSVCITVDYKLPRTVPDPLAVK 597

Qy 590 DGTSCGSKNYCRNORCVSSYL---GYDCTTDCNDRGVNCKKHCHCSASYLPDQCSVQ 646
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
598 NGSQCDIGRCVNVRECVESRIIICASAVC-SQQSCSGHGVCDNRNKHCHSGYKFPNQCIR 656

Qy 647 SLDLPGSGSI---DSGNFPFVPAIPARLPERRYRIENIYHSPMRW--PFFLIFPFFIIFCV 700
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
657 S---KGPSIFPEEDMGSI-----MERASGKTENTWLLGFLIALPILIV---696

Qy 701 LIAMVKVNRQKWRTEYSSDEQPESEPEPKG 734
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
697 -TTAIVLARKQLKWKFAKE---EFPSSSEKSEG 726
```

RESULT 8

US-10-230-338-90

; Sequence 90, Application US/10230338

; Publication No. US20030044934A1

```
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Deenoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC92
; CURRENT APPLICATION NUMBER: US/10/230,338
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 90
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-230-338-90

Query Match      38.3%; Score 1524.5; DB 14; Length 787;
Best Local Similarity 41.1%; Pred. No. 2.1e-113;
Matches 310; Conservative 134; Mismatches 259; Indels 51; Gaps 18;

Qy 1 MWVLFLLSGGLRMD-SNFDLSPVQITVPEKIRSIIEKIEGIE---SQASYKIVIEGKPYT 56
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 LWL--LLAGLCGLASRPGFQNSLLQIVPEKIQTNDSEIEYEQISYIIPIDEKLYT 61

Qy 57 VNLMOKNFLPHNPRVYSYSGTGIMKPLDDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGV 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 VHLKQRYFLADNFMILYN-QGSMNTYSSDIQTQCYQGNIEGYPDSMVTLCSTGSLRGI 120

Qy 117 LQFENVSYGIEPLESSVGFHEHYQVKKHKKADVSLYNEKDIESR---DLSPKLOQSAEPQ 173
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 LQFENVSYGIEPLESAVEFQHVLYLKLKNEQNDIAIFIDRSLSKEQPMDDNIFISEKSEPAV 180

Qy 174 D--FAKYEIMHVIVEKQLYNHMGSDTTVAQKVFQILGTNAIFVSNITIISSLELWI 231
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 PDLFPLYLEHMHVVDKTYDYGSDSMIVTNKVIEIVGLANSMTQPKVTIVLSLELWS 240

Qy 232 DENKIATTEGANELLHTFLRWKTSYLVLRPHDVAFLVREKSNYVGTGQKMCADANYA 291
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 DENKISTVGEADELLQKFLWKQSYLNLPRHDIAVLLIYMDYPRYLGAVPFGTMCITRYS 300

Qy 292 GGVLHPRTISLESIAVILAQLLSMGITYDDINKQCCSGAVCIMNPEAHFSGVKIFS 351
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 AGVALYPKEITLEAFVAVITQMLALSIGISYDDPKKQCSESTCIMNPEVQVNGVKTFS 360
```



```
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC98
; CURRENT APPLICATION NUMBER: US/10/230,414
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 90
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-230-414-90

Query Match      38.3%; Score 1524.5; DB 14; Length 787;
Best Local Similarity 41.1%; Pred. No. 2.le-113;
Matches 310; Conservative 134; Mismatches 259; Indels 51; Gaps 18;

QY 1 MWVFLSLGGLRMD-SNFDSPVQITVPEKRSIIKEGIE---SQASKYIVIEGPKYT 56
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 LWL--LLAGLCGLASRPGESLLQIVPEIKIQTNTDSSEIYEIQISYIIPIDEKLYT 61

QY 57 VNLQKQFLPHNFRVSYSGTGIMKPLDQDFQNFCHYGVIEGPKSVVMVSTCTGLGV 116
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 VHLKQRYFLADNFWIYLYN-QGSMNTYSSDIQTQCYTQGNIEGYPDMSVTLSTCSGLRGI 120

QY 117 LQFENVSYGIEPLESSYGFHVIYQVRKHKADVSLYNEKDIESR---DLSFKLQSAEPQ 173
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 LQFENVSYGIEPLESAVEFHVLYKLKNEKDIAIFDRSLKQPMDDNFISEKSEPAV 180

QY 174 D--PAKYIEMHVIKOLYHMGSDTTVAQKVPQLIGLNAIPVSNITIISSLELWI 231
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 PDLFPLYLEHVIYVDKTYLWGSDSMIVTNKVIIEIVGLANSMTQPKVTIVLSLELWS 240

QY 232 DENKIATIGRANELLHFLRWKTSYLVLRHDVAFLLAVREKSNYVCATQGMKCDANYA 291
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 DENKISIVGADELQKFLSEWKQSYLNRPHDIAILLIYMDYPRYLGAVPFGTWCTITRYS 300

QY 292 GGVVLPRTISLESVALIAQLLSLMSGITYDDINKQCCSGAVCMNPEAIHFSGVKIFS 351
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 AGVALYPKETLEAFAVIVTQMLALSIGISYDDPKKCCQSESTCINNPEVQSNVKTFS 360

QY 352 NCSFEDFAHFSIKQKSLQNLQPLDPFFQQAQVCGNAKLEAGRECDGTEQDCALIGET 411
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361 SCSLRSFQNFISNVGVKCLQKQPMQK-KSPKPVCGNRLGNEICDCGTEAQCG--PAS 417

QY 412 CCDIATCRFXAGNSCAEGPCENCLFWSKERMCRP--SFEECDLPEYKCNCGSSASCENHYV 470
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
418 CCFDRCVLKDGAKCYKGLCKCKCQILQSQGVCEKRPKAPPCDIAENCGSSPECGPDITL 477
```

```
QY 471 QTGHPGCLNQWICIDGYCMSSGDQKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNGIS-DS 529
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
478 INGLSKCKNNKFCYDGDGDCHLDARCESVFGKSGRNAPFACYEEIQSQSDRFGNGCRDRNN 537

QY 530 GYTQCEADNLCQCKLICKYVVGKFLQIPRATIIYANISGHLCTAVFASDHADSQKMWIK 589
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
538 KYVFCGWRNLICRLVCTPFRKPFHQENGDIYAFVSDSVCTVDYKLPRTVPDPLAVK 597

QY 590 DGTSCGSKNVCNORCVSSSYL---GYDCTTDKNDRGVCNNKHKCHCSASYLPPDCSVQ 646
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
598 NGSQCDIGRVNRECVESRIIKASAHVC-SQCSGHGVCDNRKNCCHCSGYKPPNQIR 656

QY 647 SLDWPGGSI---DSGNFPFPAIPARLPERRYENIYHSPKRW--PFFLFIPIFFIIFCV 700
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
657 S---KGFSIFPEEDMGSI-----MERASGKTENTWLLGFLIALPILIV--- 696
```

## RESULT 11

```
US-10-232-224-90
; Sequence 90, Application US/10232224
; Publication No. US20030065147A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC111
; CURRENT APPLICATION NUMBER: US/10/232,224
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 90
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-232-224-90
```

Query Match 38.3%; Score 1524.5; DB 14; Length 787;  
Best Local Similarity 41.1%; Pred. No. 2.le-113;







```

, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101922
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/106178
, PRIOR FILING DATE: 1998-10-28
, PRIOR APPLICATION NUMBER: 60/106248
, PRIOR FILING DATE: 1998-10-29
, PRIOR APPLICATION NUMBER: 60/106464
, PRIOR FILING DATE: 1998-10-30
, PRIOR APPLICATION NUMBER: 60/106905
, PRIOR FILING DATE: 1998-11-03
, PRIOR APPLICATION NUMBER: 60/108787
, PRIOR FILING DATE: 1998-11-17
, PRIOR APPLICATION NUMBER: 60/108801
, PRIOR FILING DATE: 1998-11-17
, PRIOR APPLICATION NUMBER: 60/108849
, PRIOR FILING DATE: 1998-11-18
, PRIOR APPLICATION NUMBER: 60/112422
, PRIOR FILING DATE: 1998-12-15
, PRIOR APPLICATION NUMBER: 60/113296
, PRIOR FILING DATE: 1998-12-22
, PRIOR APPLICATION NUMBER: 60/113605
, PRIOR FILING DATE: 1998-12-23
, PRIOR APPLICATION NUMBER: 60/113621
, PRIOR FILING DATE: 1998-12-23
, PRIOR APPLICATION NUMBER: 60/115558
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115565
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115733
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115949
, PRIOR FILING DATE: 1999-02-10
, PRIOR APPLICATION NUMBER: 60/123618
, PRIOR FILING DATE: 1999-03-10
, PRIOR APPLICATION NUMBER: 60/125259
, PRIOR FILING DATE: 1999-03-19
, PRIOR APPLICATION NUMBER: 60/125775
, PRIOR FILING DATE: 1999-03-23
, PRIOR APPLICATION NUMBER: 60/126773
, PRIOR FILING DATE: 1999-03-29
, PRIOR APPLICATION NUMBER: 60/127887
, PRIOR FILING DATE: 1999-04-05
, PRIOR APPLICATION NUMBER: 60/130232
, PRIOR FILING DATE: 1999-04-21
, PRIOR APPLICATION NUMBER: 60/131022
, PRIOR FILING DATE: 1999-04-26
, PRIOR APPLICATION NUMBER: 60/131270
, PRIOR FILING DATE: 1999-04-27
, PRIOR APPLICATION NUMBER: 60/131291
, PRIOR FILING DATE: 1999-04-27
, PRIOR APPLICATION NUMBER: 60/140650
, PRIOR FILING DATE: 1999-06-22
, PRIOR APPLICATION NUMBER: 60/141445
, PRIOR FILING DATE: 1999-04-28
, PRIOR APPLICATION NUMBER: 60/134287
, PRIOR FILING DATE: 1999-05-14
, PRIOR APPLICATION NUMBER: 60/140650
, PRIOR FILING DATE: 1999-06-22
, PRIOR APPLICATION NUMBER: 60/140723
, PRIOR FILING DATE: 1999-08-22
, PRIOR APPLICATION NUMBER: 60/141037
, PRIOR FILING DATE: 1999-06-23
, PRIOR APPLICATION NUMBER: 60/144758
, PRIOR FILING DATE: 1999-07-20
, PRIOR APPLICATION NUMBER: 60/145698
, PRIOR FILING DATE: 1999-07-26
, PRIOR APPLICATION NUMBER: 60/146222
, PRIOR FILING DATE: 1999-07-28
, PRIOR APPLICATION NUMBER: 60/146963
, PRIOR FILING DATE: 1999-08-03
, PRIOR APPLICATION NUMBER: 60/149320
, PRIOR FILING DATE: 1999-08-17
, PRIOR APPLICATION NUMBER: 60/149638
, PRIOR FILING DATE: 1999-08-17

```



;  
; PRIOR APPLICATION NUMBER: 60/096146  
; PRIOR FILING DATE: 1998-08-11  
; PRIOR APPLICATION NUMBER: 60/096791  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 60/097986  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: 60/098544  
; PRIOR FILING DATE: 1998-08-31  
; PRIOR APPLICATION NUMBER: 60/099596  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099598  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099803  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099811  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099812  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099816  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/100038  
; PRIOR FILING DATE: 1998-09-11  
; PRIOR APPLICATION NUMBER: 60/100385  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/100390  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/100627  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100848  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/100919  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/101477  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101738  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101741  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101786  
; PRIOR FILING DATE: 1998-09-25  
; PRIOR APPLICATION NUMBER: 60/101916  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101922  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/106178  
; PRIOR FILING DATE: 1998-10-28  
; PRIOR APPLICATION NUMBER: 60/106248  
; PRIOR FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: 60/106464  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 60/106905  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: 60/108787  
; PRIOR FILING DATE: 1998-11-17  
; PRIOR APPLICATION NUMBER: 60/108801  
; PRIOR FILING DATE: 1998-11-17  
; PRIOR APPLICATION NUMBER: 60/108849  
; PRIOR FILING DATE: 1998-11-18  
; PRIOR APPLICATION NUMBER: 60/112422  
; PRIOR FILING DATE: 1998-12-15  
; PRIOR APPLICATION NUMBER: 60/113296  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 60/113605  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113621  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/115558  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/115565  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/115733  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/119549

;  
; PRIOR FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 60/123618  
; PRIOR FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: 60/125259  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 60/125775  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: 60/126773  
; PRIOR FILING DATE: 1999-03-29  
; PRIOR APPLICATION NUMBER: 60/127887  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 60/130232  
; PRIOR FILING DATE: 1999-04-21  
; PRIOR APPLICATION NUMBER: 60/131022  
; PRIOR FILING DATE: 1999-04-26  
; PRIOR APPLICATION NUMBER: 60/131270  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/131291  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/131445  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: 60/134287  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/140650  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/140723  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/141037  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 60/144758  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/145698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: 60/146222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: 60/146963  
; PRIOR FILING DATE: 1999-08-03  
; PRIOR APPLICATION NUMBER: 60/149320  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/149638  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/151733  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/164418  
; PRIOR FILING DATE: 1999-11-09  
; PRIOR APPLICATION NUMBER: 60/166361  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR APPLICATION NUMBER: 60/169445  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 60/169495  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 60/169835

Query Match 38.3%; Score 1524.5; DB 14; Length 787;

Best Local Similarity 41.1%; Pred. No. 2.1e-113;

Matches 310; Conservative 134; Mismatches 259; Indels 51; Gaps 18;

Qy 1 MWVLELLSGIGGLRMD-SNPFDSLVPVOITVPEKIRSIKEGIE--SQASYKIVIEGKPYT 56

Db 4 LWL--LLAGLCGLASRPGFQNSLLQIVPEKIQTWNDSSEIEYEQISVIIPIDEKLYT 61

Qy 57 VNLQMKNFLPHNFRVYSQGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGV 116

Db 62 VHLKQRYFLADNFMVLYLN-QGSMNTYSSDIQTQCYVQGNIEGYPDSMVTLSCTSLRGI 120

Qy 117 LQFENVSYGIEPLESSVGFHEHYIYQVHKKADVSLNEXKDIESR---DLSEFKLOSAEPQQ 173

Db 121 LQFENVSYGIEPLESAVEFQHVLYLKLKNEKDNDIAIFIRSLRKEQPMDDNIFISEKSPAV 180

Qy 174 D---FAXYIEMHVIVEKOLYNHMGSDTTVVAAQKVPQLIGLTNAIFVSNITIIILSSLELWI 231

Db 181 PDLFPLYLEMHVVDKTLVDYWGSDSMIVTNKRVIEIVGLANSWFTQPKVTIVLSSLELWS 240

Qy	232	DENKIATTGCEANELLHTFLRWKTSYVLURPHDVAFLVYREKSNYVGATQGMKCDANTA	291
Db	241	DENKISTVGDEADLEKQLEWKQSYNLRPHDIAIYLIYMDPRYLGAVPPGTMCTTRY	300
Qy	292	GGVHLHPRTISLESIAVILIAQLLSMIGITVDDINKQCQSGAVCMMNPEAHFESGVKIFS	351
Db	301	AGVALYPKETITLEAFAVITQMLALSLSIGSYDDPKPKQCSSESTCIMNPEVQSGVKTFS	360
Qy	352	NGCSFEDFAHFISKQSQCLHNPRLDPFFKQOAVCGNAKUEAGBECDCGTQBDICALIBT	411
Db	361	SCSLRSFQNFISNVGVKCLQNKPQMK-KSPKPCVCGNRLEGNEICDCGTEAQCG--PAS	417
Qy	412	CDDIATCRFKAGSNCAEGPCCECNLFMSKERMCRP-SFECDLPEYCVNGSSASCSPENHVY	470
Db	418	CDDFRTCVLKQAKCVKGLCCXDCQILOSVECEPKAHPECDIAENCGSSPFCGPDITL	477
Qy	471	QTGHPCGLNQMWCIDGVCMSSGDKQCTDTFGKEVFGPSECVSHLNSKNTDVSNGCIS-DS	529
Db	478	INGLSCKNNKFCICYDGDCHDLARCESYFVGKSGSNAPFACVEEIQSOSDRFGNCGRDRNN	537
Qy	530	GYTQCEADNLQCKLICKYVGKFLLIQIPRATTIIVANISGHLCIAVEFASDHADSQKMWIK	589
Db	538	KYVFCGWRNLIQRLVCTYPTKRPQHQENGDIVIAFVDSVCITVDYKLPRTVPDPLAVK	597
Qy	590	DGTS CGSNKVCNRQCVSSYL---GYDCTTDKCDRGVCMNKKHCHCSASYLPPDCSVQ	646
Db	598	NGSQCDIGRVCVNRCEVESRIIKASAHVC-SQSCSGHGVCDSRNKCHCSGPGYPPNQCIR	656
Qy	647	SDLWPQGSII---DSGNPPFPVAIPARLPERRVNIENIYHSKPMRW--PFEFLIFPFEILFCV	700
Db	657	S---KGFSLPPEDMGSI-----MERASGKTENTWLLGLTALILIV---696	
Qy	701	LTAIMVKVNFQRKWRTEYDSDEQFSESESPKG	734
Db	697	-TTAIVLARKOLKQWFAKE---EEFPSESSEKSG	726

Search completed: January 10, 2005, 22:31:53  
Job time : 151 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:22:33 ; Search time 160 Seconds  
(without alignments)  
1645.671 Million cell updates/sec

Title: US-10-054-683-19  
Perfect score: 734  
Sequence: 1 MMVFLSLGLGLRMDNSNFD.....WRTDYSSDEQPESEPKG 734

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	734	100.0	734	2 AAR87037	Human PH3
2	734	100.0	734	6 ABJ19246	Human can
3	734	100.0	734	7 ADB75186	Prostate
4	650	88.6	651	2 AAR87034	Human PH3
5	631	86.0	735	7 ADB75184	Prostate
6	609	83.0	733	7 ADB75182	Prostate
7	15	2.0	568	4 AAE10823	Human gen
8	15	2.0	668	6 ABJ26664	Human pro
9	15	2.0	759	5 AAU97037	Human LP
10	15	2.0	787	5 AAU83636	Human PRO
11	15	2.0	787	6 ABU80783	Human PRO
12	15	2.0	787	6 ABO33749	Novel hum
13	15	2.0	787	6 ABU82092	Novel hum
14	15	2.0	787	6 ABR48487	Human LAP
15	15	2.0	787	6 ABJ72272	Human PRO
16	15	2.0	787	6 ABJ72400	Human PRO
17	15	2.0	787	6 ABO34295	Human sec
18	15	2.0	787	7 ABJ72102	Human mem
19	15	2.0	787	7 ADB83580	Novel hum
20	15	2.0	787	7 ADB80686	Novel hum
21	15	2.0	787	7 ADB73227	Novel hum
22	15	2.0	787	7 ADB78309	Novel hum
23	15	2.0	787	7 ADB84957	Human PRO
24	15	2.0	787	7 ADB78063	Novel hum
25	15	2.0	787	7 ADB87129	Human PRO

26	15	2.0	787	7 ADB84711	Human PRO
27	15	2.0	787	7 ADB83826	Novel hum
28	15	2.0	787	7 ADB72981	Novel hum
29	15	2.0	787	7 ADC36819	Human PRO
30	15	2.0	787	7 ADC21809	Human PRO
31	15	2.0	787	7 ADC49840	Novel hum
32	15	2.0	787	7 ADC49039	Novel hum
33	15	2.0	787	7 ADC49556	Novel hum
34	15	2.0	787	7 ADC47417	Novel hum
35	15	2.0	787	7 ADC47162	Novel hum
36	15	2.0	787	7 ADC78037	Novel hum
37	15	2.0	787	7 ADD06272	Novel hum
38	15	2.0	787	7 ADC77791	Novel hum
39	15	2.0	787	7 ADD50754	Novel hum
40	15	2.0	787	7 ADD51000	Novel hum
41	15	2.0	787	7 ADD50481	Human PRO
42	15	2.0	787	7 ADD50235	Human PRO
43	15	2.0	787	7 ADD51246	Novel hum
44	15	2.0	787	8 ADC48793	Novel hum
45	15	2.0	787	8 ADE20964	Novel hum

ALIGNMENTS

RESULT 1  
AAR87037  
ID AAR87037 standard; protein; 734 AA.  
XX  
AC AAR87037;  
XX  
DT 30-SEP-1996 (first entry)  
XX  
DE Human PH30 beta chain sperm protein.  
XX  
KW Human PH30 beta chain sperm protein; contraceptive;  
FEE integrin binding domain.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Region 84..734  
FT /note= "see AAR87035"  
XX  
PN WO9535118-A1.  
XX  
PD 28-DEC-1995.  
XX  
PF 06-JUN-1995; 95WO-US007295.  
XX  
PR 20-JUN-1994; 94US-00264101.  
XX  
PA (MERI ) MERCK & CO INC.  
XX  
PI Alves K, Gupta SK, Hollis GF;  
XX  
DR WPI; 1996-058212/06.  
XX  
N-PSDB; AAT07328.  
XX  
PT Human and mouse sperm protein PH30 beta chain and related DNA - useful in  
contraceptive vaccines.  
XX  
PS Example 2; Page 45-48; 85pp; English.  
XX  
CC Human PH30 beta chain sperm protein, having a FEE integrin binding  
domain, is 58.9% identical to mouse and 56.5% identical to guinea pig  
PH30 beta. The protein may be produced recombinantly and used in a  
contraceptive composition containing an effective adjuvant and an amount  
of sperm protein which is effective for the stimulation of antibodies  
which bind to sperm protein in vivo, thereby preventing or substantially  
reducing the rate of sperm-egg fusion  
XX  
SQ Sequence 734 AA;

```
Query Match      100.0%; Score 734; DB 2; Length 734;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWVLFLLSGGLGRLMDSNFDLPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNLM 60
DB 1 MWVLFLLSGGLGRLMDSNFDLPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNLM 60
QY 61 QKNFLPHNFRVYSYSGTGMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQPE 120
DB 61 QKNFLPHNFRVYSYSGTGMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQPE 120
QY 121 NVSYGIEPLESSVGFHEVIYQVHKKADVSLYNEKDIERSDLGFKLOSAPQODFAKYE 180
DB 121 NVSYGIEPLESSVGFHEVIYQVHKKADVSLYNEKDIERSDLGFKLOSAPQODFAKYE 180
QY 121 NVSYGIEPLESSVGFHEVIYQVHKKADVSLYNEKDIERSDLGFKLOSAPQODFAKYE 180
DB 121 NVSYGIEPLESSVGFHEVIYQVHKKADVSLYNEKDIERSDLGFKLOSAPQODFAKYE 180
QY 181 MHVIVEKQLYNHMGSDTTVVAQKVFQILGTLNATFVSNITIIISLLELWIDENKIATTG 240
DB 181 MHVIVEKQLYNHMGSDTTVVAQKVFQILGTLNATFVSNITIIISLLELWIDENKIATTG 240
QY 241 EANELLHTFLRWKTSYLVLRPHDVAFLLVYREKSNYVGATFQGMCDANVAGGVVLHPRT 300
DB 241 EANELLHTFLRWKTSYLVLRPHDVAFLLVYREKSNYVGATFQGMCDANVAGGVVLHPRT 300
QY 301 ISLESIAVILAOQLLSLMSGITDDINKCQCSGAVCIMNPEAIHPSGVKIFSNCSFEDFAH 360
DB 301 ISLESIAVILAOQLLSLMSGITDDINKCQCSGAVCIMNPEAIHPSGVKIFSNCSFEDFAH 360
QY 361 FISKQKSOQLHNPRLDPPFPKQAVCGNAKLEAGECDCCGTEQDCALIGTCCDIATCRF 420
DB 361 FISKQKSOQLHNPRLDPPFPKQAVCGNAKLEAGECDCCGTEQDCALIGTCCDIATCRF 420
QY 421 KAGSNCAEGPCCNCLFMSKERMCRPSFECDLPYECNCGSSASCPENHYVQTGHPCCGLNQ 480
DB 421 KAGSNCAEGPCCNCLFMSKERMCRPSFECDLPYECNCGSSASCPENHYVQTGHPCCGLNQ 480
QY 481 WICIDGVCMSGDCKOCTDTFGKEVEFGPSECVSHLNSKTDVSGNGCIGDSGVTOCEADNLQ 540
DB 481 WICIDGVCMSGDCKOCTDTFGKEVEFGPSECVSHLNSKTDVSGNGCIGDSGVTOCEADNLQ 540
QY 541 CGKLICKYVGFLLQIIPRATIIYANISGHLCIAVEFASDHADSQKMWIKDGTSCGSKVC 600
DB 541 CGKLICKYVGFLLQIIPRATIIYANISGHLCIAVEFASDHADSQKMWIKDGTSCGSKVC 600
QY 601 RNQRCVSSYLGYDCTTDKNDRGVGNKKGCHCSASYLPDSCSVQSDLPWGGSIDSGNF 660
DB 601 RNQRCVSSYLGYDCTTDKNDRGVGNKKGCHCSASYLPDSCSVQSDLPWGGSIDSGNF 660
QY 661 PPVAIPARLPERRIENIYHSKPMRWPFLLPFIPIFCVLIAIMVKNVQKWKRTEDY 720
DB 661 PPVAIPARLPERRIENIYHSKPMRWPFLLPFIPIFCVLIAIMVKNVQKWKRTEDY 720
QY 721 SSDQEPSESESPKG 734
DB 721 SSDQEPSESESPKG 734
```

## RESULT 2

```
ABJ19246
ID ABJ19246 standard; protein; 734 AA.
XX AC ABJ19246;
XX AC ABJ19246;
XX DT 28-MAR-2003 (first entry)
XX DE Human cancer/testis antigen - SEQ ID No 19.
XX KW Human; gene therapy; vaccine; cancer; cancer/testis antigen; CT antigen.
XX OS Homo sapiens.
XX PN W0200278526-A2.
```

```
XX 10-OCT-2002.
XX 29-MAR-2002; 2002WO-US009808.
XX 30-MAR-2001; 2001US-0280718P.
XX 20-APR-2001; 2001US-0285154P.
XX 05-OCT-2001; 2001US-0327432P.
XX 22-JAN-2002; 2002US-00054683.
XX (LUDW-) LUDWIG INST CANCER RES.
XX (CORR ) CORNELL RES FOUND INC.
XX Old LJ, Scanlan MJ, Chen Y;
XX WPI; 2003-040608/03.
XX N-PSDB; ABT15727.
XX Diagnosing cancer comprises contacting a biological sample isolated from
XX a subject with an agent that specifically binds to a nucleic acid
XX molecule, its expression product or fragment or an antibody that binds to
XX the product or fragment.
XX Claim 36; Page 124-126; 155pp; English.
XX The invention comprises a method for diagnosing cancer, the method
XX involves detecting the DNA or protein sequences of human cancer/testis
XX (CT) antigens that are disclosed in the invention. The method of the
XX invention is useful for detecting/diagnosing, treating and monitoring a
XX cancer or condition characterised by the expression of a human CT
XX antigen. The present amino acid sequence represents a human CT antigen of
XX the invention
XX Sequence 734 AA;
```

```
Query Match      100.0%; Score 734; DB 6; Length 734;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWVLFLLSGGLGRLMDSNFDLPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNLM 60
DB 1 MWVLFLLSGGLGRLMDSNFDLPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNLM 60
QY 61 QKNFLPHNFRVYSYSGTGMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQPE 120
DB 61 QKNFLPHNFRVYSYSGTGMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQPE 120
QY 121 NVSYGIEPLESSVGFHEVIYQVHKKADVSLYNEKDIERSDLGFKLOSAPQODFAKYE 180
DB 121 NVSYGIEPLESSVGFHEVIYQVHKKADVSLYNEKDIERSDLGFKLOSAPQODFAKYE 180
QY 181 MHVIVEKQLYNHMGSDTTVVAQKVFQILGTLNATFVSNITIIISLLELWIDENKIATTG 240
DB 181 MHVIVEKQLYNHMGSDTTVVAQKVFQILGTLNATFVSNITIIISLLELWIDENKIATTG 240
QY 241 EANELLHTFLRWKTSYLVLRPHDVAFLLVYREKSNYVGATFQGMCDANVAGGVVLHPRT 300
DB 241 EANELLHTFLRWKTSYLVLRPHDVAFLLVYREKSNYVGATFQGMCDANVAGGVVLHPRT 300
QY 301 ISLESIAVILAOQLLSLMSGITDDINKCQCSGAVCIMNPEAIHPSGVKIFSNCSFEDFAH 360
DB 301 ISLESIAVILAOQLLSLMSGITDDINKCQCSGAVCIMNPEAIHPSGVKIFSNCSFEDFAH 360
QY 361 FISKQKSOQLHNPRLDPPFPKQAVCGNAKLEAGECDCCGTEQDCALIGTCCDIATCRF 420
DB 361 FISKQKSOQLHNPRLDPPFPKQAVCGNAKLEAGECDCCGTEQDCALIGTCCDIATCRF 420
QY 421 KAGSNCAEGPCCNCLFMSKERMCRPSFECDLPYECNCGSSASCPENHYVQTGHPCCGLNQ 480
DB 421 KAGSNCAEGPCCNCLFMSKERMCRPSFECDLPYECNCGSSASCPENHYVQTGHPCCGLNQ 480
QY 481 WICIDGVCMSGDCKOCTDTFGKEVEFGPSECVSHLNSKTDVSGNGCIGDSGVTOCEADNLQ 540
DB 481 WICIDGVCMSGDCKOCTDTFGKEVEFGPSECVSHLNSKTDVSGNGCIGDSGVTOCEADNLQ 540
```

Db 481 WICIDGVCMSGDQCTDTFGKEVEFGPSECVSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540

Qy 541 CGKLICKYVGKFLQIIPRATIIYANISGHLICIAVEFASDHADSQKMWIKDGTSCGSKNVC 600

Db 541 CGKLICKYVGKFLQIIPRATIIYANISGHLICIAVEFASDHADSQKMWIKDGTSCGSKNVC 600

Qy 601 RNQRCVSSSYLGVDCTTDCNDRGVGNKKHCHCSASLYLPDSCVQSDLPFGGSIDSGNF 660

Db 601 RNQRCVSSSYLGVDCTTDCNDRGVGNKKHCHCSASLYLPDSCVQSDLPFGGSIDSGNF 660

Qy 661 PPVAIPARLPERRIYENIYHSKPMRWPFLLFIPFFIIFCVLIAMVKNFQKKWRTEDY 720

Db 661 PPVAIPARLPERRIYENIYHSKPMRWPFLLFIPFFIIFCVLIAMVKNFQKKWRTEDY 720

Qy 721 SSDEQPESESEPKG 734

Db 721 SSDEQPESESEPKG 734

RESULT 3

ID ADB75186

AC ADB75186 standard; protein; 734 AA.

XX ADB75186;

DT 04-DEC-2003 (first entry)

DE Prostate cancer marker protein.

XX Prostate; cancer; cytostatic; gene therapy; marker.

XX Homo sapiens.

XX WO2003009814-A2.

XX 06-FEB-2003.

XX 25-JUL-2002; 2002WO-US023913.

XX 25-JUL-2001; 2001US-0307982P.

PR 22-AUG-2001; 2001US-0314356P.

PR 25-SEP-2001; 2001US-0325020P.

PR 12-DEC-2001; 2001US-0341746P.

PR 05-MAR-2002; 2002US-0362158P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

PI Hoersh S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;

XX WPI; 2003-248033/24.

XX New nucleic acid molecule, useful for diagnosing or treating prostate cancer.

XX Disclosure; SEQ ID NO 10; 99pp; English.

XX The invention relates to newly discovered cancer markers associated with the cancerous state of prostate cells. Also disclosed is a method of assessing whether a patient is afflicted with prostate cancer. The method of the invention involves assessing whether a patient is afflicted with prostate cancer by comparing the level of expression of a marker in a patient sample and the normal level of expression of the marker in a control non-prostate cancer sample, where a significant increase in the level of expression of the marker in the patient sample and the normal level indicates that the patient is afflicted with prostate cancer. Nucleic acids of the invention are useful for diagnosing or treating prostate cancer, and may be useful in gene therapy. Sequences given in ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 734 AA;

Query Match 100.0%; Score 734; DB 7; Length 734;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWVLFLLSGLGGRMDSNFDLSLPVQITVPKIRSIIEGIESQASYKIVIEGKYTVNLM 60

Db 1 MWVLFLLSGLGGRMDSNFDLSLPVQITVPKIRSIIEGIESQASYKIVIEGKYTVNLM 60

Qy 61 QKMFPLNFRVYSVSGTGIMKPLDDQDFQNFCHQGYIEGYPKSVVMYSTCTGLRGVLQFE 120

Db 61 QKMFPLNFRVYSVSGTGIMKPLDDQDFQNFCHQGYIEGYPKSVVMYSTCTGLRGVLQFE 120

Qy 121 NVSYGIEPLESSVGFHVYIQVKHKADVSLYNEKDIIESRDLSPKLSAEPQDFAKYIE 180

Db 121 NVSYGIEPLESSVGFHVYIQVKHKADVSLYNEKDIIESRDLSPKLSAEPQDFAKYIE 180

Qy 181 MHVIVEKQLYNHMGSDTTVVAQVFLIGLTNAIFVSFNITIIILSSLELWIDENKIATTG 240

Db 181 MHVIVEKQLYNHMGSDTTVVAQVFLIGLTNAIFVSFNITIIILSSLELWIDENKIATTG 240

Qy 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVYGATFQGMKCDANVAGGVVLLHPT 300

Db 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVYGATFQGMKCDANVAGGVVLLHPT 300

Qy 301 ISLES LAVILAQLLSLSMGITYDDINKCQCSGAVCINMPEAIHFSGVKIFSNCSFEDFAH 360

Db 301 ISLES LAVILAQLLSLSMGITYDDINKCQCSGAVCINMPEAIHFSGVKIFSNCSFEDFAH 360

Qy 361 FISKQSQCLHNPRLDPPFPKQAVCNKLEAGEBCDCGTEQDCALIGETCCDIATCRF 420

Db 361 FISKQSQCLHNPRLDPPFPKQAVCNKLEAGEBCDCGTEQDCALIGETCCDIATCRF 420

Qy 421 KAGSNCAEGPCCNCLFMSKERNCRPSFECDDLPEYCNCGSSASCPENHYVQTGHPCGQ 480

Db 421 KAGSNCAEGPCCNCLFMSKERNCRPSFECDDLPEYCNCGSSASCPENHYVQTGHPCGQ 480

Qy 481 WICIDGVCMSGDQCTDTFGKEVEFGPSECVSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540

Db 481 WICIDGVCMSGDQCTDTFGKEVEFGPSECVSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540

Qy 541 CGKLICKYVGKFLQIIPRATIIYANISGHLICIAVEFASDHADSQKMWIKDGTSCGSKNVC 600

Db 541 CGKLICKYVGKFLQIIPRATIIYANISGHLICIAVEFASDHADSQKMWIKDGTSCGSKNVC 600

Qy 601 RNQRCVSSSYLGVDCTTDCNDRGVGNKKHCHCSASLYLPDSCVQSDLPFGGSIDSGNF 660

Db 601 RNQRCVSSSYLGVDCTTDCNDRGVGNKKHCHCSASLYLPDSCVQSDLPFGGSIDSGNF 660

Qy 661 PPVAIPARLPERRIYENIYHSKPMRWPFLLFIPFFIIFCVLIAMVKNFQKKWRTEDY 720

Db 661 PPVAIPARLPERRIYENIYHSKPMRWPFLLFIPFFIIFCVLIAMVKNFQKKWRTEDY 720

Qy 721 SSDEQPESESEPKG 734

Db 721 SSDEQPESESEPKG 734

RESULT 4

AAR87034

ID AAR87034 standard; protein; 651 AA.

XX AAR87034;

AC AAR87034;

DT 30-SEP-1996 (first entry)

DE Human PH30 beta chain sperm protein.

XX Human PH30 beta chain sperm protein; contraceptive;

KW FEE integrin binding domain.

XX Homo sapiens.

XX WO9535118-A1.  
PN XX  
PD 28-DEC-1995.  
XX  
XX 06-JUN-1995; 95WO-US007295.  
PF XX  
XX 20-JUN-1994; 94US-00264101.  
PR XX  
XX (MERI ) MERCK & CO INC.  
PA XX  
XX Alves K, Gupta SK, Hollis GF;  
PI XX  
XX WPI; 1996-038212/06.  
DR N-PSDB; AAT36700.  
XX  
XX Human and mouse sperm protein PH30 beta chain and related DNA - useful in  
PT contraceptive vaccines.  
XX  
XX Disclosure; Page 29-31; 85pp; English.  
XX  
XX Human PH30 beta chain sperm protein, having an FEE integrin binding  
CC domain, may be recombinantly produced by vector-mediated gene expression  
CC in host cells, preferably mammalian cells e.g. Chinese hamster ovary  
CC (CHO) cell culture. It may be used in a contraceptive composition  
CC containing an effective adjuvant and an amount of sperm protein which is  
CC effective for the stimulation of antibodies which bind to sperm protein  
CC in vivo, thereby preventing or substantially reducing the rate of sperm-  
CC egg fusion. The protein is also useful for identifying small molecules  
CC that disrupt sperm-egg interaction and fertilization  
XX  
XX Sequence 651 AA;  
SQ

Query Match 88.6%; Score 650; DB 2; Length 651;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 QDFQNFCHYQYIGYIPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVIYQVKH 144  
DB 2 QDFQNFCHYQYIGYIPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVIYQVKH 61  
QY 145 KKADVSLYNEKDIERSDLSPKLSAEPQDPFQAKYIEMHVIIVEKOLYNHMGSDTTVAQKV 204  
DB 62 KKADVSLYNEKDIERSDLSPKLSAEPQDPFQAKYIEMHVIIVEKOLYNHMGSDTTVAQKV 121  
QY 205 FQIGLNLAI FVSNITITIISSLELWIDENKIATTGEANELLHTFLRWKTSYLVRPHDV 264  
DB 122 FQIGLNLAI FVSNITITIISSLELWIDENKIATTGEANELLHTFLRWKTSYLVRPHDV 181  
QY 265 AFLIVYREKSNYVGATFQGMCDANYAGGVVLHPTTISLES LAVILAQLLSLMSGITYDD 324  
DB 182 AFLIVYREKSNYVGATFQGMCDANYAGGVVLHPTTISLES LAVILAQLLSLMSGITYDD 241  
QY 325 INKQCQSGAVCINMPEAIHSGVKVIFNCSPEDPFAHIFISKQSCNLNQRLDPFFKQQA 384  
DB 242 INKQCQSGAVCINMPEAIHSGVKVIFNCSPEDPFAHIFISKQSCNLNQRLDPFFKQQA 301  
QY 385 VCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERMC 444  
DB 302 VCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERMC 361  
QY 445 RPSFECDLPEYCNCGSSASCPENHYVQTGHPCGLNQWICIDGVCMSGDKQCTDTFGKEVE 504  
DB 362 RPSFECDLPEYCNCGSSASCPENHYVQTGHPCGLNQWICIDGVCMSGDKQCTDTFGKEVE 421  
QY 505 FGPSECYSHLNSKTDVSGNCGISDSGYTQCBADNLQCGKLIKCTKVGFLLQIPRATIYA 564  
DB 422 FGPSECYSHLNSKTDVSGNCGISDSGYTQCBADNLQCGKLIKCTKVGFLLQIPRATIYA 481  
QY 565 NISGHLICIAVEFASDHADSQKMWIKDGTSCGSKNVCRNQRVSSSYLGDCCTDKNDRG 624  
DB 482 NISGHLICIAVEFASDHADSQKMWIKDGTSCGSKNVCRNQRVSSSYLGDCCTDKNDRG 541

QY 625 VCNNKXCHCSASYLPPDCSVQSDLMFPGSGIDSGNPPPPVAIPARLPERRYIENIYHMKPM 684  
DB 542 VCNNKXCHCSASYLPPDCSVQSDLMFPGSGIDSGNPPPPVAIPARLPERRYIENIYHMKPM 601  
QY 685 RWPFFLPIPPFIIFCVLLIIMVKNVQKWKRTEDYSSDQPESESEPKG 734  
DB 602 RWPFFLPIPPFIIFCVLLIIMVKNVQKWKRTEDYSSDQPESESEPKG 651  
RESULT 5  
ADB75184  
ID ADB75184 standard; protein; 735 AA.  
XX  
XX ADB75184;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Prostate cancer marker protein.  
XX  
KW Prostate; cancer; cytostatic; gene therapy; marker.  
XX  
OS Homo sapiens.  
XX  
PN WO2003009814-A2.  
XX  
PD 06-FEB-2003.  
XX  
PF 25-JUL-2002; 2002WO-US023913.  
XX  
PR 25-JUL-2001; 2001US-0307982P.  
PR 22-AUG-2001; 2001US-0314356P.  
PR 25-SEP-2001; 2001US-0325020P.  
PR 12-DEC-2001; 2001US-0341746P.  
PR 05-MAR-2002; 2002US-0362158P.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;  
PI Hoersh S, Kamatkar S, Woney AM, Glatt K, Zhao X, Anderson D;  
XX  
DR WPI; 2003-248033/24.  
XX  
XX New nucleic acid molecule, useful for diagnosing or treating prostate  
PT cancer.  
XX  
PS Claim 4; SEQ ID NO 8; 99pp; English.  
XX  
CC The invention relates to newly discovered cancer markers associated with  
CC the cancerous state of prostate cells. Also disclosed is a method of  
CC assessing whether a patient is afflicted with prostate cancer. The method  
CC of the invention involves assessing whether a patient is afflicted with  
CC prostate cancer by comparing the level of expression of a marker in a  
CC patient sample and the normal level of expression of the marker in a  
CC control non-prostate cancer sample, where a significant increase in the  
CC level of expression of the marker in the patient sample and the normal  
CC level indicates that the patient is afflicted with prostate cancer.  
CC Nucleic acids of the invention are useful for diagnosing or treating  
CC prostate cancer, and may be useful in gene therapy. Sequences given in  
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 735 AA;

Query Match 86.0%; Score 631; DB 7; Length 735;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 731; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 VLFSLGLGLRMDSNFDSLPLVQITVPEKIRSIIEKIESQASYKIVIEGKPYTNLMQK 62  
DB 4 VLFSLGLGLRMDSNFDSLPLVQITVPEKIRSIIEKIESQASYKIVIEGKPYTNLMQK 63



QY 63 NPLPHNPRVYSYSGTGIMKPLDQDFQNFCHYQYIEGPKSVVMVSTCTGLRGVLQFENV 122  
Db 64 NPLPHNPRVYSYSGTGIMKPLDQDFQNFCHYQYIEGPKSVVMVSTCTGLRGVLQFENV 123  
QY 123 SYGIEPLESSVGFHEVYIQVKKKADVSLYNEKDIESRDLSPKLOSAEPQDDPAKYIEMH 182  
Db 124 SYGIEPLESSVGFHEVYIQVKKKADVSLYNEKDIESRDLSPKLOSAEPQDDPAKYIEMH 183  
QY 183 VIVEKQLYNHMGSDTTTVAQVQFQIGLTNAIFVSNFTIILSSLELWIDENKIATPTEA 242  
Db 184 VIVEKQLYNHMGSDTTTVAQVQFQIGLTNAIFVSNFTIILSSLELWIDENKIATPTEA 243  
QY 243 NEILLHFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFQKMCDDANVAGVVLHPTIS 302  
Db 244 NEILLHFLRWKTSYLVLRPHDVAFLVYREKSNVVGATFQKMCDDANVAGVVLHPTIS 303  
QY 303 LESLAVILAQLLSMIGITVDDINKKOCGAVCIIMNPEALHFGVKIFSNCSFEDPAHFI 362  
Db 304 LESLAVILAQLLSMIGITVDDINKKOCGAVCIIMNPEALHFGVKIFSNCSFEDPAHFI 363  
QY 363 SKQSKQCLHNPRLDPPFKQQAQVCGNAKLEAGEECDCGTQDCALIGETCCDIATCRFKA 422  
Db 364 SKQSKQCLHNPRLDPPFKQQAQVCGNAKLEAGEECDCGTQDCALIGETCCDIATCRFKA 423  
QY 423 GSNCAEGPCENCLFMSKERMCRPSFECDLPRYCNCGSSASCENHYVQTHGFCGLNQWI 482  
Db 424 GSNCAEGPCENCLFMSKERMCRPSFECDLPRYCNCGSSASCENHYVQTHGFCGLNQWI 483  
QY 483 CIDGVCMSGDKQCTDTDFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCG 542  
Db 484 CIDGVCMSGDKQCTDTDFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCG 543  
QY 543 KLICKYVKGFLQIPRATIIYANISGHLCIAVEFADSHADSQKMWIKDGTSCGNKVCYN 602  
Db 544 KLICKYVKGFLQIPRATIIYANISGHLCIAVEFADSHADSQKMWIKDGTSCGNKVCYN 603  
QY 603 QRCVSSSYLGYDCTTDCNDRGVNKKHCHCSASYLPDDCSVQSDLWPGGSDSGNFPF 662  
Db 604 QRCVSSSYLGYDCTTDCNDRGVNKKHCHCSASYLPDDCSVQSDLWPGGSDSGNFPF 663  
QY 663 VAIPARLPERRYENIYHSPMRWPFLEFPFIFCVLIAIMVKNVFORKKWRTEDYSS 722  
Db 664 VAIPARLPERRYENIYHSPMRWPFLEFPFIFCVLIAIMVKNVFORKKWRTEDYSS 723  
QY 723 DEQSESEPKG 734  
Db 724 DEQSESEPKG 735  
RESULT 6  
ID ADB75182  
XX ADB75182 standard; protein; 753 AA.  
AC ADB75182;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Prostate cancer marker protein.  
XX  
XX Prostate; cancer; cytostatic; gene therapy; marker.  
XX  
OS Homo sapiens.  
XX  
XX WO2003009814-A2.  
XX  
PD 06-FEB-2003.  
XX  
XX 25-JUL-2002; 2002WO-US023913.  
XX  
XX 25-JUL-2001; 2001US-0307982P.  
PR 22-AUG-2001; 2001US-0314356P.  
PR 25-SEP-2001; 2001US-0325020P.  
PR 12-DEC-2001; 2001US-0341746P.

PR 05-MAR-2002; 2002US-0362158P.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;  
PI Hoerh S, Kamatkar S, Woneey AM, Glatk K, Zhao X, Anderson D;  
XX WPI; 2003-248033/24.  
DR  
XX New nucleic acid molecule, useful for diagnosing or treating prostate  
PT cancer.  
XX  
PS Claim 4; SEQ ID NO 6; 99pp; English.  
XX  
XX The invention relates to newly discovered cancer markers associated with  
CC the cancerous state of prostate cells. Also disclosed is a method of  
CC assessing whether a patient is afflicted with prostate cancer. The method  
CC of the invention involves assessing whether a patient is afflicted with  
CC prostate cancer by comparing the level of expression of a marker in a  
CC patient sample and the normal level of expression of the marker in a  
CC control non-prostate cancer sample, where a significant increase in the  
CC level of expression of the marker in the patient sample and the normal  
CC level indicates that the patient is afflicted with prostate cancer.  
CC Nucleic acids of the invention are useful for diagnosing or treating  
CC prostate cancer, and may be useful in gene therapy. Sequences given in  
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 753 AA;  
Query Match 83.0%; Score 609; DB 7; Length 753;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 709; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 25 QITVPEKIRSIIEGIESQASQYKIVIEGKPYTNLMQKFLPHNFRVYSYSGTGIMKPLD 84  
Db 44 QITVPEKIRSIIEGIESQASQYKIVIEGKPYTNLMQKFLPHNFRVYSYSGTGIMKPLD 103  
QY 85 QDFQNFCHYQYIEGPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVYQVXH 144  
Db 104 QDFQNFCHYQYIEGPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVYQVXH 163  
QY 145 KKADVSLYNEKDIESRDLSPKLOSAEPQDDPAKYIEMHIVKQLYNHMGSDTTTVAQKV 204  
Db 164 KKADVSLYNEKDIESRDLSPKLOSAEPQDDPAKYIEMHIVKQLYNHMGSDTTTVAQKV 223  
QY 205 FQIGLTNAIFVSNFTIILSSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 264  
Db 224 FQIGLTNAIFVSNFTIILSSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 283  
QY 265 AFLVYREKSNVVGATFQKMCDDANVAGVVLHPTISLESILAQLLSMIGITVDD 324  
Db 284 AFLVYREKSNVVGATFQKMCDDANVAGVVLHPTISLESILAQLLSMIGITVDD 343  
QY 325 INKQCQSGAVCIIMNPEALHFGVKIFSNCSFEDPAHFIKQSKQCLHNPRLDPPFKQQA 384  
Db 344 INKQCQSGAVCIIMNPEALHFGVKIFSNCSFEDPAHFIKQSKQCLHNPRLDPPFKQQA 403  
QY 385 VCGNAKLEAGEECDCGTQDCALIGETCCDIATCRPKAGNCNCAEGPCENCLFMSKERM 444  
Db 404 VCGNAKLEAGEECDCGTQDCALIGETCCDIATCRPKAGNCNCAEGPCENCLFMSKERM 463  
QY 445 RPSFECDLPRYCNCGSSASCENHYVQTHGFCGLNQWICIDGVCMSGDKQCTDTDFGKEVE 504  
Db 464 RPSFECDLPRYCNCGSSASCENHYVQTHGFCGLNQWICIDGVCMSGDKQCTDTDFGKEVE 523  
QY 505 FGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLI CKYVKGFLQIPRATIIYA 564  
Db 524 FGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLI CKYVKGFLQIPRATIIYA 583  
QY 565 NISGHLCIAVEFADSHADSQKMWIKDGTSCGNKVCYNQRCVSSSYLGYDCTTDCNDRG 624

Db	584	NISGHLCIAVEFASDHADSKWMIKOGTSCGSKNVCQRCVSSYLGXDCTDKNDRG	643
Qy	625	VCNKKKHCHCSASYLPDSCVQSDLPFGGSDGNSFPFPAIPARLPERRIENIYHSPM	684
Db	644	VCNKKKHCHCSASYLPDSCVQSDLPFGGSDGNSFPFPAIPARLPERRIENIYHSPM	703
Qy	685	RWPFELPFPFIIFCVLIAMVKVNFORKKWRTEYSSDQPESESEPKG	734
Db	704	RWPFELPFPFIIFCVLIAMVKVNFORKKWRTEYSSDQPESESEPKG	753
RESULT 7			
AAE10823			
ID	AAE10823	standard; protein; 568 AA.	
AC	AAE10823;		
XX			
DT	18-DEC-2001	(first entry)	
XX			
DE	Human gene 2 encoded ADAM protein HOEKN33, SEQ ID NO:7.		
XX			
KW	Human; ADAM protein; a disintegrin and metalloprotease domain; cancer;		
KW	adamalysin; angiogenic disorder; chronic inflammatory disorder; ARDS;		
KW	rheumatoid arthritis; adult respiratory distress syndrome; asthma;		
KW	Crohn's disease; atherosclerosis; macular degeneration; psoriasis;		
KW	diabetic retinopathy; Alzheimer's disease; tissue remodeling;		
KW	haemorrhagic disorder; infertility; bone related disorder; stroke;		
KW	autoimmune disorder; haematopoietic disorder; infectious disease;		
KW	HIV-induced dementia; Human immunodeficiency virus; arrhythmia; epilepsy;		
KW	pulmonary disorder; central nervous system disorder;		
KW	respiratory disorder; obesity; cachexia; wasting disease; anorexia;		
KW	food additive; food preservative; gene therapy.		
XX			
OS	Homo sapiens.		
XX			
Key		Location/Qualifiers	
PH	Region	6..17	
FT		/note="Epitope"	
FT	Region	72..77	
FT		/note="Epitope"	
FT	Region	86..93	
FT		/note="Epitope"	
FT	Region	100..105	
FT		/note="Epitope"	
FT	Region	188..193	
FT		/note="Epitope"	
FT	Region	248..254	
FT		/note="Epitope"	
FT	Region	268..274	
FT		/note="Epitope"	
FT	Region	284..291	
FT		/note="Epitope"	
FT	Region	302..307	
FT		/note="Epitope"	
FT	Region	340..358	
FT		/note="Epitope"	
FT	Region	377..388	
FT		/note="Epitope"	
FT	Region	405..410	
FT		/note="Epitope"	
FT	Region	418..423	
FT		/note="Epitope"	
FT	Region	456..472	
FT		/note="Epitope"	
FT	Region	495..501	
FT		/note="Epitope"	
FT	Region	533..548	
FT		/note="Epitope"	
FT	Region	552..568	
FT		/note="Epitope"	
XX			

PN	WO200166557-A1.		
XX			
PD	13-SEP-2001.		
XX			
PF	22-FEB-2001; 2001WO-US005497.		
XX			
PR	03-MAR-2000; 2000US-0187937P.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Shi Y, Ruben SM;		
XX			
DR	WPI; 2001-596823/67.		
DR	N-PSDB; AAD18168.		
XX			
PT	Novel proteins comprising contain a disintegrin and metalloprotease		
PT	domain, referred as ADAM and ADAM polynucleotides for diagnosing,		
PT	preventing and treating cancer, inflammatory, reproductive,		
PT	cardiovascular disorders.		
XX			
PS	Claim 11; Page 274-276; 280pp; English.		
XX			
CC	AAD18167-AD18170 represent cDNAs corresponding to human ADAM protein		
CC	(proteins which contain a disintegrin and metalloprotease domain, also		
CC	called adamalysin) genes, and AAE10822-AAE10825 represent the proteins		
CC	they encode. ADAM proteins and their corresponding genes are useful for		
CC	preventing, treating or ameliorating a medical condition in a mammal.		
CC	Sequences of the invention and their antibodies are useful for diagnosis,		
CC	and treatment of diseases related to angiogenic disorders such as cancer,		
CC	and cancer metastasis, chronic inflammatory disorders such as rheumatoid		
CC	arthritis, hepatitis, nephritis, Crohn's disease, asthma, adult		
CC	respiratory distress syndrome (ARDS), atherosclerosis, macular		
CC	degeneration, psoriasis, haemorrhagic disorders, infertility, disorders		
CC	remodeling the skeletal system, reproductive system, bone related		
CC	disorders, autoimmune disorders, haematopoietic disorders, IGF-mediated		
CC	allergic reactions, organ transplant rejections, graft-versus-host		
CC	disease, infectious diseases, HIV-induced dementia, arrhythmias, high		
CC	blood pressure, muscular contractile dysfunction, pace-maker dysfunction,		
CC	disorder of proper neurotransmitter release, epilepsy, stroke, neural,		
CC	gastrointestinal (e.g. inflammatory bowel disease), cardiovascular (e.g.		
CC	myocarditis), renal, pulmonary, proliferative, respiratory (e.g. allergy)		
CC	disorders and/or cancerous diseases or conditions such as gastric, lung,		
CC	ovarian, bladder, liver and breast), central nervous system disorders		
CC	(e.g. prion disease) and/or hormone secretion disorders. ADAM genes,		
CC	proteins, agonists or antagonist prevents skin aging due to sunburn,		
CC	modulate the differentiation or proliferation of embryonic stem cells,		
CC	modulate mammalian characteristics such as body height, weight, hair		
CC	colour, skin, modulate mammalian metabolism, biornhms and are useful		
CC	for treating obesity, cachexia, wasting disease, anorexia and bulimia.		
CC	They are also useful as food additive or preservative to increase or		
CC	decrease storage capabilities, fat content, lipid, protein or other		
CC	nutritional components. Polynucleotides of the invention are useful in		
CC	gene therapy. The present sequence represents a human ADAM protein of the		
CC	invention		
XX			
SQ	Sequence 568 AA;		
		Query Match	2.0%; Score 15; DB 4; Length 568;
		Best Local Similarity	100.0%; Pred. No. 1.5e-05;
		Matches	15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	117 LQFENVSYGIEPLES 131		
Db	47 LQFENVSYGIEPLES 61		
RESULT 8			
ABJ26664			
ID	ABJ26664	standard; protein; 668 AA.	
XX			
AC	ABJ26664;		
XX			

DT XX 01-MAY-2003 (first entry)  
DE XX Human protein modification + maintenance molecule protein SEQ ID No 18.  
XX XX Cystostatic; antiarteriosclerotic; anti-HIV; antiallergic; anticonvulsant;  
KW KW cerebroprotective; antiparkinsonian; nootropic; antiinflammatory;  
KW KW antitumor; hepatotropic; gynecological; antibacterial; virucide;  
KW KW protozoacide; antiparasitic; cell proliferative disease; PMOD;  
KW KW protein modification and maintenance molecule; immunogenic fragment;  
KW KW cancer; autoimmune; inflammatory disease; neurological disorder;  
KW KW gastrointestinal; developmental; vesicle trafficking disorder; infection;  
KW KW protein-protein interaction; drug-target interaction;  
KW KW gene expression profile; human.  
XX XX Homo sapiens.  
OS XX WO2003000844-A2.  
XX XX 03-JAN-2003.  
XX XX 18-JUN-2002; 2002WO-US019360.  
XX XX 22-JUN-2001; 2001US-0300508P.  
XX XX 06-JUL-2001; 2001US-0303445P.  
XX XX 13-JUL-2001; 2001US-0305405P.  
XX XX 09-AUG-2001; 2001US-0311442P.  
XX XX 24-AUG-2001; 2001US-0314821P.  
XX XX 29-AUG-2001; 2001US-0315992P.  
XX XX 03-MAY-2002; 2002US-0378205P.  
XX XX (INCY-) INCYTE GENOMICS INC.  
XX XX Gandhi AR, Kable AE, Swarnakar A, Hafalia AJA, Tran B, Duggan BM;  
PI PI Warren BA, Ison CH, Honeckell CD, Nguyen DB, Lu DAM, Lee EA, Yue H;  
PI PI Forsythe IJ, Barroso I, Rankumar J, Griffin JA, Li JX, Yang J;  
PI PI Thangavelu K, Gietzen KJ, Ding L, Baughn MR, Borowsky ML, Yao MG;  
PI PI Wallia NK, Mason PM, Gururajan R, Lee S, Becha SD, Tran UK;  
PI PI Elliott VS, Luo W, Sprague WW, Tang YT, Lu Y, Zebbarjadian Y;  
XX XX WPI; 2003-184039/18.  
DR DR N-PSDB; ABT23217.  
XX XX New isolated human PMOD polypeptide and polynucleotide, useful for  
PT PT diagnosing, treating and preventing diseases or conditions associated  
PT PT with the aberrant PMOD expression, e.g. cancer, AIDS, atherosclerosis and  
PT PT infections.  
XX XX Claim 73; Page 193-194; 225pp; English.  
XX XX The invention relates to an isolated polypeptide comprising: any of 28  
CC CC sequences of 48-1256 amino acids; a natural amino acid sequence at least  
CC CC 90% identical to the 28 amino acid sequences, 94% identical to a sequence  
CC CC of 703 or 267 amino acids, 96% identical to a sequence of 414 amino  
CC CC acids, or 9% identical to a sequence of 242 amino acids, all given in  
CC CC the specification; or a biologically active or immunogenic fragment of  
CC CC the isolated polypeptide. The polypeptides and polynucleotides are useful  
CC CC in diagnosing, treating and preventing diseases or conditions associated  
CC CC with the decreased expression of protein modification and maintenance  
CC CC molecules (PMOD), such as cell proliferative diseases (e.g. cancer,  
CC CC atherosclerosis), autoimmune/inflammatory diseases (e.g. AIDS,  
CC CC allergies), neurological disorders (e.g. stroke, Parkinson's disease,  
CC CC epilepsy), gastrointestinal (e.g. ulcer, cirrhosis), reproductive (e.g.  
CC CC endometriosis), developmental, vesicle trafficking disorders, and  
CC CC infections (e.g. bacterial, viral, parasitic, protozoal). These are also  
CC CC useful in assessing the effects of exogenous compounds on the expression  
CC CC of nucleic acid and amino acid sequences of PMOD. The PMOD or its  
CC CC fragments are useful in screening compounds for effectiveness as agonist  
CC CC or antagonist of the polypeptides, or in altering the expression of the  
CC CC target polynucleotide and compounds that specifically bind to or modulate  
CC CC the activity of the polypeptide. The microarray is useful in monitoring  
CC CC or measuring protein-protein interactions, drug-target interactions, and  
CC CC gene expression profiles. This sequence represents a human PMOD protein  
CC CC of the invention

XX SQ Sequence 668 AA;  
XX XX Query Match 2.0%; Score 15; DB 6; Length 668;  
XX XX Best Local Similarity 100.0%; Pred. No. 1.7e-05; Indels 0; Gaps 0;  
XX XX Matches 15; Conservative 0; Mismatches 0;  
Oy 117 LQFENVSYGIEPLES 131  
Db 121 LQFENVSYGIEPLES 135  
|||||  
RESULT 9  
AAU97037  
ID AAU97037 standard; protein; 759 AA.  
XX AC AAU97037;  
XX DT 13-AUG-2002 (first entry)  
XX DE Human LP protein LP102.  
XX KW Human; LP102; inflammatory disorder; diabetes; bone disease;  
KW KW cardiovascular disease; male reproductive system disease; osteoporosis;  
KW KW Paget's disease; myeloma; Alzheimer's disease; contraceptive;  
KW KW liver cancer; growth factor-mediated disease; anaphylaxis; coagulation;  
KW KW sepsis; skeletal muscle dystrophy; asthma; breast cancer.  
XX OS Homo sapiens.  
XX PN WO200232939-A2.  
XX PD 25-APR-2002.  
XX PF 10-OCT-2001; 2001WO-US027759.  
XX PR 19-OCT-2000; 2000US-0241813P.  
XX PA (ELIL ) LILLY & CO ELI.  
XX PI Lu D, Song HY, Su EW, Wang H;  
XX XX WPI; 2002-454591/49.  
DR DR N-PSDB; ABKS1494.  
XX XX New secreted human LP polypeptides or polynucleotides, useful for  
PT PT treating mammals suffering from conditions associated with aberrant  
PT PT levels of an LP polypeptide, e.g. cancers, osteoporosis, Paget's disease  
PT PT or Alzheimer's disease.  
XX XX Claim 8; Page 131-134; 148pp; English.  
XX CC The invention relates to isolated human polypeptides designated LP102,  
CC CC LP187, LP190 and LP241, and the polynucleotides encoding them. The LP  
CC CC polypeptide or the LP polypeptide agonist is useful for treating a mammal  
CC CC suffering from a disease, condition or disorder associated with aberrant  
CC CC levels of an LP polypeptide. The LP polypeptides or polynucleotides, or  
CC CC the antibodies are useful for treating, preventing or diagnosing cancers,  
CC CC inflammatory disorders, bone diseases or cardiovascular diseases. In  
CC CC particular, the LP102 polypeptides, polynucleotides or antibodies are  
CC CC useful for diagnosing, preventing or treating male reproductive system  
CC CC diseases, osteoporosis, Paget's disease, metastatic or myeloma associated  
CC CC bone diseases or Alzheimer's disease. These are also useful as  
CC CC contraceptive agents. The LP187 polypeptides, polynucleotides or  
CC CC antibodies are particularly useful for diagnosing, preventing or treating  
CC CC liver cancer, as well as other growth factor-mediated diseases and  
CC CC conditions. The LP190 polypeptides, polynucleotides or antibodies are  
CC CC useful for diagnosing, preventing or treating asthma, anaphylaxis, and  
CC CC diseases related to coagulation or sepsis. LP241 polypeptides,  
CC CC polynucleotides or antibodies are useful especially for diagnosing,  
CC CC preventing or treating skeletal muscle dystrophy, breast cancer or  
CC CC diabetes. The LP polynucleotides are also useful for constructing DNA  
CC CC vectors that may be employed in medicine, or for preparing the proteins.

CC The present sequence represents the amino acid sequence of LP102

XX  
SQ Sequence 759 AA;

Query Match 2.0%; Score 15; DB 5; Length 759;

Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFENVSYGIEPLES 131

Db 121 LQFENVSYGIEPLES 135

RESULT 10

AAU83636 ID AAU83636 standard; protein; 787 AA.

XX AAU83636;

XX 08-MAY-2002 (first entry)

DE Human PRO protein, Seq ID No 90.

XX Human; secreted protein; PRO; tumour; lung cancer; colon cancer;

KW breast cancer; prostate tumour; rectal tumour; liver tumour;

KW pericyte cell proliferation; chondrocyte cell proliferation;

KW tumour necrosis factor-alpha.

XX Homo sapiens.

XX WO200208288-A2.

XX 31-JAN-2002.

XX 29-JUN-2001; 2001WO-US021066.

XX 20-JUL-2000; 2000US-0219556P.

XX 25-JUL-2000; 2000US-0220585P.

XX 25-JUL-2000; 2000US-0220605P.

XX 25-JUL-2000; 2000US-0220607P.

XX 25-JUL-2000; 2000US-0220624P.

XX 25-JUL-2000; 2000US-0220638P.

XX 25-JUL-2000; 2000US-0220664P.

XX 25-JUL-2000; 2000US-0220666P.

XX 26-JUL-2000; 2000US-0220893P.

XX 28-JUL-2000; 2000WO-US020710.

XX 01-AUG-2000; 2000US-0222425P.

XX 22-AUG-2000; 2000US-02271133P.

XX 23-AUG-2000; 2000WO-US023522.

XX 24-AUG-2000; 2000WO-US023328.

XX 10-NOV-2000; 2000WO-US030873.

XX 28-NOV-2000; 2000US-0253646P.

XX 01-DEC-2000; 2000WO-US032878.

XX 20-DEC-2000; 2000US-00747259.

XX 20-DEC-2000; 2000WO-US034956.

XX 28-FEB-2001; 2001WO-US006520.

XX 01-MAR-2001; 2001WO-US006566.

XX 22-MAR-2001; 2001US-00816744.

XX 10-MAY-2001; 2001US-00854208.

XX 25-MAY-2001; 2001US-00854280.

XX 25-MAY-2001; 2001WO-US017092.

XX (GETH ) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2002-172001/22.

XX N-PSDB; ABK33580.

XX One hundred and twenty two nucleic acids encoding PRO polypeptides,

PT useful for treating a PRO related disorder and for diagnosing tumors such

PT as lung cancer, colon cancer, breast tumor, prostate tumor, rectal tumor

PT or liver tumor.

XX Claim 11; Fig 90; 359pp; English.

XX The invention relates to one hundred and twenty two nucleic acids  
CC encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides  
CC encode human secreted proteins. The PRO nucleic acids, polypeptides,  
CC agonists and antagonists are useful for treating a PRO related disorder.  
CC The PRO polypeptides are useful for diagnosing tumours, especially lung  
CC cancer, colon cancer, breast tumor, prostate tumor, rectal tumor or  
CC liver tumor. The PRO polypeptides are useful for stimulating the  
CC proliferation of, or gene expression, in pericyte cells, for stimulating  
CC the proliferation or differentiation of chondrocyte cells, for  
CC stimulating the release of tumour necrosis factor-alpha from human blood,  
CC for stimulating or inhibiting the proliferation of normal human dermal  
CC fibroblast cells. The PRO polypeptide may also be used as molecular  
CC weight markers and for tissue typing. The PRO nucleic acids have  
CC applications in molecular biology, including use as hybridisation probes,  
CC and in chromosome and gene mapping. AAU83592-AAU83713 represent human PRO  
CC protein sequences of the invention

XX SQ Sequence 787 AA;

Query Match 2.0%; Score 15; DB 5; Length 787;

Best Local Similarity 100.0%; Pred. No. 1.9e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFENVSYGIEPLES 131

Db 121 LQFENVSYGIEPLES 135

RESULT 11

ABU80783 ID ABU80783 standard; protein; 787 AA.

XX AC ABU80783;

XX 23-JUN-2003 (first entry)

XX Human PRO polypeptide #45.

XX Human; PRO polypeptide; secreted and transmembrane protein;

KW anti-PRO antibody; diagnostic assay; gene expression; tumour; cytostatic.

XX Homo sapiens.

XX US2003036635-A1.

XX 20-FEB-2003.

XX 28-AUG-2002; 2002US-00230163.

XX 25-JUL-2000; 2000US-0220638P.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH ) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-342045/32.

XX N-PSDB; ACA66885.

XX One hundred and twenty two nucleic acids encoding PRO polypeptides,

PT useful for the manufacture of a medicament for diagnosing or treating

PT tumor.

XX Claim 11; Fig 90; 314pp; English.

XX The present invention relates to the isolation of novel human PRO

CC polypeptides, and the polynucleotide sequences encoding them. The PRO  
 CC polypeptides are secreted and transmembrane proteins. The PRO  
 CC polypeptides and polynucleotides are useful for preparing a medicament  
 CC useful in the diagnosis and treatment of tumours. Anti-PRO antibodies are  
 CC useful in diagnostic assays for PRO, by detecting its expression in  
 CC specific cells, tissues or serum, and for affinity purification of PRO  
 CC from recombinant cell culture or natural sources. AB080739-AB080860  
 CC represent the human PRO polypeptides of the invention. Note: The sequence  
 CC data for this patent was obtained in electronic format directly from the  
 CC USPTO web site at seqdata.uspto.gov/psipdIDEntry.html  
 XX  
 SQ Sequence 787 AA;

Query Match 2.0%; Score 15; DB 6; Length 787;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFENVSYGIEPLES 131  
 |||||  
 Db 121 LQFENVSYGIEPLES 135

RESULT 12  
 AB033749  
 ID AB033749 standard; protein; 787 AA.

AC AB033749;

XX 17-SEP-2003 (first entry)

DE Novel human secreted and transmembrane protein PRO21340.

XX Human; secreted and transmembrane protein; PRO; cytostatic;  
 KW antiarthritic; osteopathic; gene therapy; TNF-Agonist-Alpha;  
 KW chondrocyte stimulator; pericyte stimulator; fibroblast modulator;  
 KW pharmaceutical; diagnostic; biosensor; bioindicator; tumour; lung tumour;  
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;  
 KW liver tumour; bone disorder; cartilage disorder; sports injury;  
 KW arthritis; wound.

XX Homo sapiens.

XX US2003045687-A1.

XX 06-MAR-2003.

XX 12-AUG-2002; 2002US-00218631.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH ) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-512315/48.

XX N-PSDB; ACD68637.

XX New genes, and its encoded secreted and transmembrane polypeptides,  
 PT useful for stimulating Tumor Necrosis Factor alpha, or chondrocyte or  
 PT pericyte proliferation, especially for treating lung tumors, arthritis or  
 PT wounds in a mammal.

XX Claim 11; Fig 90; 314pp; English.

XX The invention describes an isolated nucleic acid molecule comprising a  
 CC sequence with at least 80% identity to: (a) a nucleotide encoding any of  
 CC 122 PRO (secreted and transmembrane) polypeptides whose sequences are  
 CC fully defined in the specification; or (b) any of 122 nucleotide  
 CC sequences having e.g. 4834, 2504 or 1759 bp fully defined in the  
 CC specification; or the full length coding sequence of any of these 122

CC nucleotide sequences. The PRO polypeptides or polynucleotides are useful  
 CC as pharmaceuticals, diagnostics, biosensors or bioreactors. These are  
 CC particularly useful for detecting tumours (e.g. lung tumour, colon  
 CC tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour)  
 CC in a mammal, for stimulating the release of TNF-alpha from human blood,  
 CC for stimulating the proliferation or differentiation of chondrocyte  
 CC cells, for stimulating proliferation of pericyte cells, or for modulating  
 CC normal human dermal fibroblast proliferation. The PRO nucleic acid or  
 CC polypeptide is also useful for treating tumours or various bone and/or  
 CC cartilage disorders (e.g. sports injuries or arthritis), or wounds. The  
 CC PRO polypeptides are useful in drug screening, particularly as targets  
 CC for therapeutic intervention in these diseases, and in the diagnostic  
 CC determination of the presence of these diseases. The PRO polypeptides are  
 CC also useful as molecular weight markers, or for chromosome  
 CC identification. The PRO genes are useful as hybridisation probes, or for  
 CC screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may  
 CC also be used in gene therapy, particularly for replacing a defective  
 CC gene. This is the amino acid sequence of a novel human secreted and  
 CC transmembrane PRO polypeptide  
 XX  
 SQ Sequence 787 AA;

Query Match 2.0%; Score 15; DB 6; Length 787;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFENVSYGIEPLES 131  
 |||||  
 Db 121 LQFENVSYGIEPLES 135

RESULT 13  
 AB082092

ID AB082092 standard; protein; 787 AA.

AC AB082092;

XX 25-JUN-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO21340.

XX Human; secreted and transmembrane protein; PRO; cardiant; cytostatic;  
 KW antiangiogenic; hypotensive; vulnenry; antiarteriosclerotic;  
 KW gene therapy; cardiovascular disorder; endothelial disorder;  
 KW angiogenic disorder; cardiac hypertrophy; trauma; cancer;  
 KW age-related macular degeneration; atherosclerosis; hypertension;  
 KW arterial restenosis; rheumatoid arthritis; angina; myocardial infarction;  
 KW thrombophlebitis; lymphangitis; tumour angiogenesis; breast carcinoma;  
 KW liver carcinoma; wound healing; chromosome mapping; gene mapping.

XX Homo sapiens.

XX US2003088063-A1.

XX 08-MAY-2003.

XX 12-AUG-2002; 2002US-00219003.

XX 25-JUL-2000; 2000US-0220664P.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH ) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-393229/37.

XX N-PSDB; ACA68541.

XX One hundred and eighty seven nucleic acids encoding PRO polypeptides,  
 PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial

PT infarction), endothelial or angiogenic disorders in a mammal.  
XX Claim 11; Fig 90; 314pp; English.  
XX  
CC The invention describes one hundred and eighty seven nucleic acids  
CC encoding novel human secreted and transmembrane (PRO) polypeptides. The  
CC PRO nucleic acids, polypeptides, agonists and antagonists are useful for  
CC treating or diagnosing a cardiovascular, endothelial or angiogenic  
CC disorder in a mammal, e.g. cardiac hypertrophy, trauma, cancer, age-  
CC related macular degeneration, atherosclerosis, hypertension, arterial  
CC restenosis, rheumatoid arthritis, angina, myocardial infarctions,  
CC thrombophlebitis, lymphangitis, tumour angiogenesis (such as breast  
CC carcinoma and liver carcinoma) and wound healing. The PRO nucleic acids  
CC have applications in molecular biology, including use as hybridisation  
CC probes, and in chromosome and gene mapping. This is the amino acid  
CC sequence of a novel human secreted and transmembrane PRO polypeptide  
SQ Sequence 787 AA;  
  
Query Match 2.0%; Score 15; DB 6; Length 787;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 117 LQFENVSYGIEPLES 131  
Db 121 LQFENVSYGIEPLES 135  
|||||  
  
RESULT 14  
ABR48487  
ID ABR48487 standard; protein; 787 AA.  
XX  
AC ABR48487;  
XX  
DT 13-JUN-2003 (first entry)  
XX  
DE Human LAP cell surface protein.  
XX  
KW Human; GENSET; therapeutic; therapy.  
XX  
OS Homo sapiens.  
XX  
XN WO200294864-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 06-AUG-2001; 2001WO-IB001715.  
XX  
PR 25-MAY-2001; 2001US-0293574P.  
XX  
PR 15-JUN-2001; 2001US-0298698P.  
XX  
PR 29-JUN-2001; 2001US-0302277P.  
XX  
PR 13-JUL-2001; 2001US-0305456P.  
XX  
PA (GEST ) GENSET.  
XX  
PI Bejanin S, Tanaka H;  
XX  
DR WPI; 2003-129412/12.  
XX  
DR N-PSDB; ACC51094.  
XX  
PT New GENSET polynucleotides and polypeptides, useful for preparing a  
PT composition for treating GENSET-related disorders and as reagents in  
PT assays to quantitatively determined levels of GENSET expression in  
PT biological samples.  
XX  
PS Claim 2; Page 462-463; 505pp; English.  
XX  
CC The present invention relates to novel human GENSET coding sequences  
CC (ACC51060-ACC51115) and proteins (ABR48453-ABR48508). The GENSET  
CC sequences are useful for preparing a composition for treating GENSET-  
CC related disorders. They can also be used as markers for tissues in which  
CC the corresponding protein is preferentially expressed, as molecular  
CC weight markers on Southern gels, as chromosome markers or tags to

CC identify chromosomes, and as reagents in assays to quantitatively  
CC determined levels of GENSET expression in biological samples  
XX  
SQ Sequence 787 AA;  
  
Query Match 2.0%; Score 15; DB 6; Length 787;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 117 LQFENVSYGIEPLES 131  
Db 121 LQFENVSYGIEPLES 135  
|||||  
  
RESULT 15  
ABJ72272  
ID ABJ72272 standard; protein; 787 AA.  
XX  
AC ABJ72272;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human PRO21340 protein.  
XX  
KW PRO; proliferation; pericyte cell; TNF-alpha; blood; chondrocyte;  
KW differentiation; dermal fibroblast; tumour; gene therapy; cytostatic.  
XX  
OS Homo sapiens.  
XX  
XN US2003050448-A1.  
XX  
PD 13-MAR-2003.  
XX  
PF 28-AUG-2002; 2002US-00230414.  
XX  
PR 01-JUN-2001; 2001WO-US017800.  
XX  
PR 29-JUN-2001; 2001WO-US021066.  
XX  
PR 09-APR-2002; 2002US-00119480.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Desnoyers L, Gerriteen ME, Goddard A, Godowski PJ;  
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;  
XX  
XX N-PSDB; ABT44270.  
XX  
PT New nucleic acid encoding for a PRO protein, useful for the manufacture  
PT of a medicament for diagnosing or treating tumors or for measuring or  
PT detecting expression of an associated gene.  
XX  
PS Claim 11; Fig 90; 315pp; English.  
XX  
CC The invention relates to a novel isolated nucleic acid encoding a fully  
CC defined PRO polypeptide. The molecules of the invention may be useful for  
CC stimulating proliferation or gene expression in pericyte cells or the  
CC release of TNF-alpha from human blood. Other possible uses include the  
CC stimulation or inhibition of chondrocyte proliferation or  
CC proliferation, the stimulation of human dermal fibroblast cell  
CC proliferation and the detection of the presence of a tumour within a  
CC mammal. Furthermore, the nucleic acid may be useful for the manufacture  
CC of a medicament for diagnosing or treating a tumour within a mammal or  
CC for measuring or detecting the expression of an associated gene, as well  
CC as during gene therapy. The current sequence is that of the human PRO  
CC protein of the invention.  
XX  
SQ Sequence 787 AA;  
  
Query Match 2.0%; Score 15; DB 6; Length 787;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 117 LQFENVSYGIEPLES 131

Db           |||||||  
          121 LQFENVSYGIEPLES 135

Search completed: January 10, 2005, 22:34:38  
Job time : 162 secs

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:31:59 ; Search time 148 Seconds  
(without alignments)  
1788.239 Million cell updates/sec

Title: US-10-054-683-19  
Perfect score: 734  
Sequence: 1 MWVFLSLGLGRLMDSNF.....WRTEYSSDEQPSSEPKG 734

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1603904 seqs, 360571292 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1603904

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : Published Applications AA.\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	734	100.0	734	14	US-10-054-683-19 Sequence 19, Appl
2	734	100.0	734	14	US-10-205-823-10 Sequence 10, Appl
3	734	100.0	734	17	US-10-473-603-19 Sequence 19, Appl
4	631	86.0	735	14	US-10-205-823-8 Sequence 8, Appl
5	609	83.0	733	14	US-10-205-823-6 Sequence 6, Appl
6	15	2.0	588	9	US-09-945-676-7 Sequence 7, Appl
7	15	2.0	759	16	US-10-381-820A-2 Sequence 2, Appl
8	15	2.0	787	10	US-09-992-600A-70 Sequence 70, Appl
9	15	2.0	787	10	US-09-924-340-70 Sequence 70, Appl
10	15	2.0	787	10	US-09-992-095B-70 Sequence 70, Appl
11	15	2.0	787	10	US-09-999-570-70 Sequence 70, Appl
12	15	2.0	787	14	US-10-227-884-90 Sequence 90, Appl
13	15	2.0	787	14	US-10-230-163-90 Sequence 90, Appl

14	15	2.0	787	14	US-10-230-338-90	Sequence 90, Appl
15	15	2.0	787	14	US-10-218-631-90	Sequence 90, Appl
16	15	2.0	787	14	US-10-230-414-90	Sequence 90, Appl
17	15	2.0	787	14	US-10-232-224-90	Sequence 90, Appl
18	15	2.0	787	14	US-10-216-159A-90	Sequence 90, Appl
19	15	2.0	787	14	US-10-218-849-90	Sequence 90, Appl
20	15	2.0	787	14	US-10-227-873-90	Sequence 90, Appl
21	15	2.0	787	14	US-10-227-883-90	Sequence 90, Appl
22	15	2.0	787	14	US-10-219-076-90	Sequence 90, Appl
23	15	2.0	787	14	US-10-230-434-90	Sequence 90, Appl
24	15	2.0	787	14	US-10-219-003-90	Sequence 90, Appl
25	15	2.0	787	14	US-10-219-075-90	Sequence 90, Appl
26	15	2.0	787	14	US-10-219-464-90	Sequence 90, Appl
27	15	2.0	787	14	US-10-219-466-90	Sequence 90, Appl
28	15	2.0	787	14	US-10-219-479-90	Sequence 90, Appl
29	15	2.0	787	14	US-10-219-481-90	Sequence 90, Appl
30	15	2.0	787	14	US-10-230-260-90	Sequence 90, Appl
31	15	2.0	787	14	US-10-232-231-90	Sequence 90, Appl
32	15	2.0	787	14	US-10-232-233-90	Sequence 90, Appl
33	15	2.0	787	14	US-10-000-489-70	Sequence 70, Appl
34	15	2.0	787	14	US-10-216-165-90	Sequence 90, Appl
35	15	2.0	787	14	US-10-218-956-90	Sequence 90, Appl
36	15	2.0	787	14	US-10-219-468-90	Sequence 90, Appl
37	15	2.0	787	14	US-10-219-478-90	Sequence 90, Appl
38	15	2.0	787	14	US-10-219-536-90	Sequence 90, Appl
39	15	2.0	787	14	US-10-000-986-70	Sequence 70, Appl
40	15	2.0	787	14	US-10-233-205-90	Sequence 90, Appl
41	15	2.0	787	14	US-10-219-072-90	Sequence 90, Appl
42	15	2.0	787	14	US-10-219-470-90	Sequence 90, Appl
43	15	2.0	787	14	US-10-219-474-90	Sequence 90, Appl
44	15	2.0	787	14	US-10-219-524-90	Sequence 90, Appl
45	15	2.0	787	14	US-10-219-528-90	Sequence 90, Appl

ALIGNMENTS

RESULT 1  
US-10-054-683-19  
; Sequence 19, Application US/10054683  
; Publication No. US20030044813A1  
; GENERAL INFORMATION:  
; APPLICANT: Old, Lloyd J.  
; APPLICANT: Scanlan, Matthew J.  
; APPLICANT: Chen, Yao-Tseeng  
; TITLE OF INVENTION: Cancer-Testis Antigens  
; FILE REFERENCE: LQ461/7125 (JRV)  
; CURRENT APPLICATION NUMBER: US/10/054,683  
; PRIOR FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: US 60/280,718  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 60/285,154  
; PRIOR FILING DATE: 2001-04-20  
; PRIOR APPLICATION NUMBER: US 60/327,432  
; PRIOR FILING DATE: 2001-10-05  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 734  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-054-683-19  
Query Match 100.0%; Score 734; DB 14; Length 734;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MWVFLSLGLGRLMDSNFVQITVPEKIRSIKEGIESQASVKIVIEGKPYTNLM 60  
Db 1 MWVFLSLGLGRLMDSNFVQITVPEKIRSIKEGIESQASVKIVIEGKPYTNLM 60  
QY 61 QKNFLPHNFRVYSYSGTGMKPLDQDFNCFHYGVKSVVMVSTCTGLRGLQFE 120

Db 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFE 120  
QY 121 NVSYGIEPLESSVGFHVIYQVXHKKADVSLYNEKDIERSDLGFKLSABPQODFAKYIE 180  
Db 121 NVSYGIEPLESSVGFHVIYQVXHKKADVSLYNEKDIERSDLGFKLSABPQODFAKYIE 180  
QY 181 MHVIVEKQLYNHMGSDTTVAQVQFQGLIGLITNAIFVSNFTIILSSLELWIDENKIATTG 240  
Db 181 MHVIVEKQLYNHMGSDTTVAQVQFQGLIGLITNAIFVSNFTIILSSLELWIDENKIATTG 240  
QY 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVYVATFQGMKCDANVAGGVWLHPRT 300  
Db 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVYVATFQGMKCDANVAGGVWLHPRT 300  
QY 301 ISLES LAVILAQLLSLSMGITDINKKQCSGAVCINMPEAIHFGSVKIFSNCSFEDFAH 360  
Db 301 ISLES LAVILAQLLSLSMGITDINKKQCSGAVCINMPEAIHFGSVKIFSNCSFEDFAH 360  
QY 361 FISKQKSQCLHNPRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420  
Db 361 FISKQKSQCLHNPRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420  
QY 421 KAGSNCAEGPCCCNCLFMSKERMCRPSFECDLPEYCNCGSSASC PENHYVQTGHPCLNQ 480  
Db 421 KAGSNCAEGPCCCNCLFMSKERMCRPSFECDLPEYCNCGSSASC PENHYVQTGHPCLNQ 480  
QY 481 WICIDGVCMSGDQKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540  
Db 481 WICIDGVCMSGDQKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540  
QY 541 CGKLIKCYVGFLLQIPRATIIYANISGHLCTIAVEFASDHADSKQMWIKDGTSCGSKNVC 600  
Db 541 CGKLIKCYVGFLLQIPRATIIYANISGHLCTIAVEFASDHADSKQMWIKDGTSCGSKNVC 600  
QY 601 RNORCVSSSYLVGYDCTTDKNDRGVCNNKHKCHCSASYLPDPCSVQSDLPWGSISDGNF 660  
Db 601 RNORCVSSSYLVGYDCTTDKNDRGVCNNKHKCHCSASYLPDPCSVQSDLPWGSISDGNF 660  
QY 661 PPVAIPARLPERRYENIYHSPKMRWPFLLFIPFFIIFCVLIAIMVKVNFQKWKRTEDY 720  
Db 661 PPVAIPARLPERRYENIYHSPKMRWPFLLFIPFFIIFCVLIAIMVKVNFQKWKRTEDY 720  
QY 721 SSDEQPESESEPKG 734  
Db 721 SSDEQPESESEPKG 734

## RESULT 2

US-10-205-823-10  
; Sequence 10, Application US/10205823  
; Publication No. US20030108963A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Gorbacheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Konsey, Angela M.  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Anderson, Dustin  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; THERAPY OF PROSTATE CANCER  
; FILE REFERENCE: MRI-044  
; CURRENT APPLICATION NUMBER: US/10/205,823  
; CURRENT FILING DATE: 2002-07-25  
; PRIOR APPLICATION NUMBER: 60/307,982  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: 60/314,356  
; PRIOR FILING DATE: 2001-08-22

; PRIOR APPLICATION NUMBER: 60/325,020  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: 60/341,746  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/362,158  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 455  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 734  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-205-823-10  
Query Match 100.0%; Score 734; DB 14; Length 734;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVLFLSLGLGLRMDNSFSLPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNLM 60  
Db 1 MVLFLSLGLGLRMDNSFSLPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNLM 60  
QY 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFE 120  
Db 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFE 120  
QY 121 NVSYGIEPLESSVGFHVIYQVXHKKADVSLYNEKDIERSDLGFKLSABPQODFAKYIE 180  
Db 121 NVSYGIEPLESSVGFHVIYQVXHKKADVSLYNEKDIERSDLGFKLSABPQODFAKYIE 180  
QY 181 MHVIVEKQLYNHMGSDTTVAQVQFQGLIGLITNAIFVSNFTIILSSLELWIDENKIATTG 240  
Db 181 MHVIVEKQLYNHMGSDTTVAQVQFQGLIGLITNAIFVSNFTIILSSLELWIDENKIATTG 240  
QY 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVYVATFQGMKCDANVAGGVWLHPRT 300  
Db 241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVYVATFQGMKCDANVAGGVWLHPRT 300  
QY 301 ISLES LAVILAQLLSLSMGITDINKKQCSGAVCINMPEAIHFGSVKIFSNCSFEDFAH 360  
Db 301 ISLES LAVILAQLLSLSMGITDINKKQCSGAVCINMPEAIHFGSVKIFSNCSFEDFAH 360  
QY 361 FISKQKSQCLHNPRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420  
Db 361 FISKQKSQCLHNPRLDPPFFKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRF 420  
QY 421 KAGSNCAEGPCCCNCLFMSKERMCRPSFECDLPEYCNCGSSASC PENHYVQTGHPCLNQ 480  
Db 421 KAGSNCAEGPCCCNCLFMSKERMCRPSFECDLPEYCNCGSSASC PENHYVQTGHPCLNQ 480  
QY 481 WICIDGVCMSGDQKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540  
Db 481 WICIDGVCMSGDQKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540  
QY 541 CGKLIKCYVGFLLQIPRATIIYANISGHLCTIAVEFASDHADSKQMWIKDGTSCGSKNVC 600  
Db 541 CGKLIKCYVGFLLQIPRATIIYANISGHLCTIAVEFASDHADSKQMWIKDGTSCGSKNVC 600  
QY 601 RNORCVSSSYLVGYDCTTDKNDRGVCNNKHKCHCSASYLPDPCSVQSDLPWGSISDGNF 660  
Db 601 RNORCVSSSYLVGYDCTTDKNDRGVCNNKHKCHCSASYLPDPCSVQSDLPWGSISDGNF 660  
QY 661 PPVAIPARLPERRYENIYHSPKMRWPFLLFIPFFIIFCVLIAIMVKVNFQKWKRTEDY 720  
Db 661 PPVAIPARLPERRYENIYHSPKMRWPFLLFIPFFIIFCVLIAIMVKVNFQKWKRTEDY 720  
QY 721 SSDEQPESESEPKG 734  
Db 721 SSDEQPESESEPKG 734

## RESULT 3

US-10-473-603-19



```
Db 244 NELLHTFLRWKTSYLVLRPHDVAFLVYREKSNVGTATFGKMCDAANYAGGVVLRPTIS 303
QY 303 LESLAVILAQLLSLSMGITYDDINKQCAGVACIMNPEAIHFGVKIFNCSFEDFAHPT 362
Db 304 LESLAVILAQLLSLSMGITYDDINKQCAGVACIMNPEAIHFGVKIFNCSFEDFAHPT 363
QY 363 SKOKSCLHNQPLDPPFFKQAVCGNAKLEAGEBCDGTEDCALIGETCCDIATCRFKA 422
Db 364 SKOKSCLHNQPLDPPFFKQAVCGNAKLEAGEBCDGTEDCALIGETCCDIATCRFKA 423
QY 423 GSNCAAGPCCENCLFMSKERMCRPSFEBCDLPYCNCGSSASCENHYVOTGHPGCLNQWI 482
Db 424 GSNCAAGPCCENCLFMSKERMCRPSFEBCDLPYCNCGSSASCENHYVOTGHPGCLNQWI 483
QY 483 CIDGVCMSGDGKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCG 542
Db 484 CIDGVCMSGDGKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCG 543
QY 543 KLIKCVGKELLOIPRATIIYANISGHLCTAVFASDHADSKQMWIKDGTSCGSKNVCRN 602
Db 544 KLIKCVGKELLOIPRATIIYANISGHLCTAVFASDHADSKQMWIKDGTSCGSKNVCRN 603
QY 603 QRCVSSSYLYGYDCTTDKCNDRGVCNNKKHCHCSASYLPPDCSVQSDLWPGGSDSGNFPP 662
Db 604 QRCVSSSYLYGYDCTTDKCNDRGVCNNKKHCHCSASYLPPDCSVQSDLWPGGSDSGNFPP 663
QY 663 VAIPARLPERRYENIYHSPMRWPPFLFPFIIFCVLIAIMVKVNFQKKWRTEDYSS 722
Db 664 VAIPARLPERRYENIYHSPMRWPPFLFPFIIFCVLIAIMVKVNFQKKWRTEDYSS 723
QY 723 DEQPESESEPKG 734
Db 724 DEQPESESEPKG 735
```

RESULT 5

```
US-10-054-683-19
; Sequence 6, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 753
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-205-823-6
Query Match 83.0%; Score 609; DB 14; Length 753;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 709; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 25 QITVPEKRSIIKEGIESQASYKIVIEGKPYTVNLQKNFLPHNFRVYSYSGTGIMKPLD 84
Db 44 QITVPEKRSIIKEGIESQASYKIVIEGKPYTVNLQKNFLPHNFRVYSYSGTGIMKPLD 103
QY 85 QDFQNFCHYQYIEGYPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFPHVYQVXH 144
Db 104 QDFQNFCHYQYIEGYPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFPHVYQVXH 163
QY 145 KKADVSLYNEKDIERSDLSFKLOSAPQODFAKYIEMHVIVVEKOLYNHMGSDTTVVAAKV 204
Db 164 KKADVSLYNEKDIERSDLSFKLOSAPQODFAKYIEMHVIVVEKOLYNHMGSDTTVVAAKV 223
QY 205 FOLIGLTNAIFVSNITITILSSLELWIDENKIATTCGEANELLHTFLRWKTSYLVLRPHDV 264
Db 224 FOLIGLTNAIFVSNITITILSSLELWIDENKIATTCGEANELLHTFLRWKTSYLVLRPHDV 283
QY 265 AFLVYREKSNVGTATFQGMKCDANYAGGVVLRPHPTISLESILAVILAQLLSLSMGITYDD 324
Db 284 AFLVYREKSNVGTATFQGMKCDANYAGGVVLRPHPTISLESILAVILAQLLSLSMGITYDD 343
QY 325 INKCCSGGAVCIMNPEAIHFGVKIFNSCSFEDFAHFKQSKQCLHNQPLDPPFKQQA 384
Db 344 INKCCSGGAVCIMNPEAIHFGVKIFNSCSFEDFAHFKQSKQCLHNQPLDPPFKQQA 403
QY 385 VCGNAKLEAGEBCDGTEDCALIGETCCDIATCRFKAGSNCAEGPCCENCLFMSKERM 444
Db 404 VCGNAKLEAGEBCDGTEDCALIGETCCDIATCRFKAGSNCAEGPCCENCLFMSKERM 463
QY 445 RPSFECDLPYCNCGSSASCENHYVOTGHPGCLNQWICIDGVCMSGDGKQCTDTFGKEVE 504
Db 464 RPSFECDLPYCNCGSSASCENHYVOTGHPGCLNQWICIDGVCMSGDGKQCTDTFGKEVE 523
QY 505 FGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLIKCYVGKFLLOIPRATIIYA 564
Db 524 FGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLIKCYVGKFLLOIPRATIIYA 583
QY 565 NISGHLCTIAVEFASDHADSKQMWIKDGTSCGSKNVCNORCVSSSYLGVDCTTDKCNDRG 624
Db 584 NISGHLCTIAVEFASDHADSKQMWIKDGTSCGSKNVCNORCVSSSYLGVDCTTDKCNDRG 643
QY 625 VCNKKKHCHCSASYLPPDCSVQSDLWPGGSDSGNFPFVAIPARLPERRYENIYHSPKM 684
Db 644 VCNKKKHCHCSASYLPPDCSVQSDLWPGGSDSGNFPFVAIPARLPERRYENIYHSPKM 703
QY 685 RWPFFLPFIPFFIIFCVLIAIMVKVNFQKKWRTEDYSSDEQPESESEPKG 734
Db 704 RWPFFLPFIPFFIIFCVLIAIMVKVNFQKKWRTEDYSSDEQPESESEPKG 753
```

RESULT 6

```
US-09-945-676-7
; Sequence 7, Application US/09945676
; Patent No. US20020077465A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT051P1
; CURRENT APPLICATION NUMBER: US/09/945,676
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: PCT/US01/05497
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/187,937
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
```

```
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-676-7

Query Match          2.0%; Score 15; DB 9; Length 568;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFENVSYGIEPLES 131
Db 47 LQFENVSYGIEPLES 61

RESULT 7
US-10-381-820A-2
; Sequence 2; Application US/10381820A
; Publication No. US20040142333A1
; GENERAL INFORMATION:
; APPLICANT: Dethun, Lu
; APPLICANT: Song, Ho Yeong
; APPLICANT: Su, Eric Wen
; APPLICANT: Wang, He
; TITLE OF INVENTION: Novel Secreted Proteins and Their Uses
; FILE REFERENCE: X-13974
; CURRENT APPLICATION NUMBER: US/10/381,820A
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 2
; LENGTH: 759
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-381-820A-2

Query Match          2.0%; Score 15; DB 16; Length 759;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFENVSYGIEPLES 131
Db 121 LQFENVSYGIEPLES 135

RESULT 8
US-09-992-600A-70
; Sequence 70; Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Jpatent
; SEQ ID NO 70
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Homo sapiens

Query Match          2.0%; Score 15; DB 10; Length 787;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFENVSYGIEPLES 131
Db 121 LQFENVSYGIEPLES 135

RESULT 9
US-09-924-340-70
; Sequence 70; Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 70
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SIGNAL
; LOCATION: 1..16
US-09-924-340-70

Query Match          2.0%; Score 15; DB 10; Length 787;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFENVSYGIEPLES 131
Db 121 LQFENVSYGIEPLES 135

RESULT 10
US-09-992-095B-70
; Sequence 70; Application US/09992095B
; Publication No. US20030157485A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US5.DIV
; CURRENT APPLICATION NUMBER: US/09/992,095B
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
```

;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/293,574  
;; PRIOR FILING DATE: 2001-05-25  
;; NUMBER OF SEQ ID NOS: 112  
;; SOFTWARE: JPatent  
;; SEQ ID NO 70  
;; LENGTH: 787  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: SIGNAL  
;; LOCATION: 1..16  
US-09-992-095B-70

Query Match 2.0%; Score 15; DB 10; Length 787;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 LQFENVSYGIEPLES 131  
Db 121 LQFENVSYGIEPLES 135

RESULT 11  
US-09-999-570-70  
;; Sequence 70, Application US/09999570  
;; Publication No. US20030170628A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Benjamin, Stephane  
;; APPLICANT: Tanaka, Hiroaki  
;; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
;; FILE REFERENCE: G-091US08DIV  
;; CURRENT APPLICATION NUMBER: US/09/999,570  
;; CURRENT FILING DATE: 2001-06-14  
;; PRIOR APPLICATION NUMBER: US 09/924,340  
;; PRIOR FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
;; PRIOR FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: US 60/305,456  
;; PRIOR FILING DATE: 2001-07-13  
;; PRIOR APPLICATION NUMBER: US 60/302,277  
;; PRIOR FILING DATE: 2001-06-29  
;; PRIOR APPLICATION NUMBER: US 60/298,698  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/293,574  
;; PRIOR FILING DATE: 2001-05-25  
;; NUMBER OF SEQ ID NOS: 112  
;; SOFTWARE: JPatent  
;; SEQ ID NO 70  
;; LENGTH: 787  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: SIGNAL  
;; LOCATION: 1..16  
US-09-999-570-70

Query Match 2.0%; Score 15; DB 10; Length 787;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 LQFENVSYGIEPLES 131  
Db 121 LQFENVSYGIEPLES 135

RESULT 12  
US-10-227-884-90  
;; Sequence 90, Application US/10227884  
;; Publication No. US20030027988A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Desnoyers, Luc

;; APPLICANT: Gerritsen, Mary  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Smith, Victoria  
;; APPLICANT: Stephan, Jean-Philippe F.  
;; APPLICANT: Watanabe, Colin L.  
;; APPLICANT: Wood, William I.  
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3530PLC79  
;; CURRENT APPLICATION NUMBER: US/10/227,884  
;; CURRENT FILING DATE: 2002-08-26  
;; PRIOR APPLICATION NUMBER: 10/119,480  
;; PRIOR FILING DATE: 2002-04-09  
;; PRIOR APPLICATION NUMBER: 60/059113  
;; PRIOR FILING DATE: 1997-09-17  
;; PRIOR APPLICATION NUMBER: 60/062287  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/063549  
;; PRIOR FILING DATE: 1997-10-28  
;; PRIOR APPLICATION NUMBER: 60/064103  
;; PRIOR FILING DATE: 1997-10-31  
;; PRIOR APPLICATION NUMBER: 60/069873  
;; PRIOR FILING DATE: 1997-12-17  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079656  
;; PRIOR FILING DATE: 1998-03-26  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/081819  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081955  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082804  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086392  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089905  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090691  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/095302  
;; PRIOR FILING DATE: 1998-08-04  
;; PRIOR APPLICATION NUMBER: 60/095318  
;; PRIOR FILING DATE: 1998-08-04  
;; PRIOR APPLICATION NUMBER: 60/095916  
;; PRIOR FILING DATE: 1998-08-10  
;; PRIOR APPLICATION NUMBER: 60/096146  
;; PRIOR FILING DATE: 1998-08-11  
;; PRIOR APPLICATION NUMBER: 60/096791

```

RESULT 13
US-10-230-163-90
; Sequence 90, Application US/10230163
; Publication No. US2003003635A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.

```

APPLICANT: Watanabe, Colin L.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3530P1C96  
CURRENT APPLICATION NUMBER: US/10/230,163  
CURRENT FILING DATE: 2002-08-28  
PRIOR APPLICATION NUMBER: 10/119,480  
PRIOR FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063549  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/064103  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/069873  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086392  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089905  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090691  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/095302  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/095318  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/095916  
PRIOR FILING DATE: 1998-08-10  
PRIOR APPLICATION NUMBER: 60/096146  
PRIOR FILING DATE: 1998-08-11  
PRIOR APPLICATION NUMBER: 60/096791  
PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 60/097986  
PRIOR FILING DATE: 1998-08-26  
PRIOR APPLICATION NUMBER: 60/098544  
PRIOR FILING DATE: 1998-08-31  
PRIOR APPLICATION NUMBER: 60/099596  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 60/099598  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 60/099803  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099811  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099812  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099816  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/100038  
PRIOR FILING DATE: 1998-09-11  
PRIOR APPLICATION NUMBER: 60/100385  
PRIOR FILING DATE: 1998-09-15  
PRIOR APPLICATION NUMBER: 60/100390  
PRIOR FILING DATE: 1998-09-15  
PRIOR APPLICATION NUMBER: 60/100627  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: 60/100848  
PRIOR FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: 60/100919  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/101477  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101738  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101741  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101786  
PRIOR FILING DATE: 1998-09-25  
PRIOR APPLICATION NUMBER: 60/101916  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101922  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/106178  
PRIOR FILING DATE: 1998-10-28  
PRIOR APPLICATION NUMBER: 60/106248  
PRIOR FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: 60/106464  
PRIOR FILING DATE: 1998-10-30  
PRIOR APPLICATION NUMBER: 60/106905  
PRIOR FILING DATE: 1998-11-03  
PRIOR APPLICATION NUMBER: 60/108787  
PRIOR FILING DATE: 1998-11-17  
PRIOR APPLICATION NUMBER: 60/108801  
PRIOR FILING DATE: 1998-11-17  
PRIOR APPLICATION NUMBER: 60/108849  
PRIOR FILING DATE: 1998-11-18  
PRIOR APPLICATION NUMBER: 60/112422  
PRIOR FILING DATE: 1998-12-15  
PRIOR APPLICATION NUMBER: 60/113296  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/113605  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/113621  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/115558  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115565  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115733  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/119549  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 60/123618  
PRIOR FILING DATE: 1999-03-10  
PRIOR APPLICATION NUMBER: 60/125259  
PRIOR FILING DATE: 1999-03-19  
PRIOR APPLICATION NUMBER: 60/125775  
PRIOR FILING DATE: 1999-03-23  
PRIOR APPLICATION NUMBER: 60/126773  
PRIOR FILING DATE: 1999-03-29  
PRIOR APPLICATION NUMBER: 60/127887



; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 60/130232  
; PRIOR FILING DATE: 1999-04-21  
; PRIOR APPLICATION NUMBER: 60/131022  
; PRIOR FILING DATE: 1999-04-26  
; PRIOR APPLICATION NUMBER: 60/131270  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/131291  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/131445  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: 60/134287  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/140650  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/140723  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/141037  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 60/144758  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/145698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: 60/146222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: 60/146963  
; PRIOR FILING DATE: 1999-08-03  
; PRIOR APPLICATION NUMBER: 60/149320  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/149638  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/151733  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: 60/164418  
; PRIOR FILING DATE: 1999-11-09  
; PRIOR APPLICATION NUMBER: 60/166361  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR APPLICATION NUMBER: 60/169445  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 60/169495  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 60/169835

Query Match 2.0%; Score 15; DB 14; Length 787;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 LQFENVSYGIEPLES 131  
|||||  
Db 121 LQFENVSYGIEPLES 135

RESULT 14  
US-10-230-338-90  
; Sequence 90, Application US/10230338  
; Publication No. US20030044934A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Gerritsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: F3530P1C92  
; CURRENT APPLICATION NUMBER: US/10/230,338  
; CURRENT FILING DATE: 2002-08-28

; PRIOR APPLICATION NUMBER: 10/119,480  
; PRIOR FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062287  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063549  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/069873  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 246  
; SEQ ID NO 90  
; LENGTH: 787  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-230-338-90

Query Match 2.0%; Score 15; DB 14; Length 787;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 LQFENVSYGIEPLES 131  
|||||  
Db 121 LQFENVSYGIEPLES 135

RESULT 15  
US-10-218-631-90  
; Sequence 90, Application US/10218631  
; Publication No. US20030045687A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Gerritsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: F3530P1C14  
; CURRENT APPLICATION NUMBER: US/10/218,631  
; CURRENT FILING DATE: 2002-08-12  
; PRIOR APPLICATION NUMBER: 10/119,480  
; PRIOR FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062287  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063549  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/069873  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:29:24 ; Search time 40 Seconds  
(without alignments)  
1216.935 Million cell updates/sec

Title: US-10-054-683-19  
Perfect score: 734  
Sequence: 1 MWVFLSLGGLGRMDSNFD.....WRTDYSDQPESEPKG 734

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 478139 seqs, 66318000 residues

Word size : 0

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*

2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*

3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*

4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pcp.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	734	100.0	734	2	US-08-765-243-8
2	734	100.0	734	5	PCT-US95-07295-8
3	650	88.6	651	1	US-08-264-101-2
4	650	88.6	651	2	US-08-765-243-2
5	650	88.6	651	5	PCT-US95-07295-2
6	15	2.0	787	4	US-10-000-489-70
7	12	1.6	48	5	PCT-US93-05640-34
8	12	1.6	50	5	PCT-US93-05640-16
9	12	1.6	50	5	PCT-US93-05640-25
10	12	1.6	735	2	US-08-765-243-6
11	12	1.6	735	5	PCT-US95-07295-6
12	10	1.4	40	5	PCT-US93-05640-12
13	10	1.4	40	5	PCT-US93-05640-23
14	10	1.4	68	3	US-08-993-165-16
15	10	1.4	68	4	US-09-540-448-16
16	10	1.4	68	4	US-09-243-640-14
17	10	1.4	68	4	US-08-929-847-16
18	10	1.4	68	4	US-09-813-484-16
19	10	1.4	71	1	US-07-985-674-8
20	10	1.4	71	5	PCT-US93-09523-8
21	10	1.4	73	6	5318899-4
22	9	1.2	23	6	5318899-75
23	9	1.2	27	5	PCT-US93-05640-14
24	9	1.2	71	1	US-07-985-674-10
25	9	1.2	71	4	US-09-460-295B-7
26	9	1.2	71	5	PCT-US93-09523-10
27	9	1.2	72	1	US-07-623-611-1

28	9	1.2	72	5	PCT-US91-09108-1	Sequence 1, Appli
29	9	1.2	72	6	5182260-21	Patent No. 5182260
30	9	1.2	72	6	5318899-7	Patent No. 5318899
31	9	1.2	72	6	5318899-14	Patent No. 5318899
32	9	1.2	73	4	US-09-776-268A-1	Sequence 1, Appli
33	9	1.2	73	4	US-09-460-295B-10	Sequence 10, Appli
34	9	1.2	73	6	5182260-2	Patent No. 5182260
35	9	1.2	73	6	5318899-2	Patent No. 5318899
36	9	1.2	73	6	5318899-3	Patent No. 5318899
37	9	1.2	73	6	5318899-6	Patent No. 5318899
38	9	1.2	73	6	5318899-10	Patent No. 5318899
39	9	1.2	95	6	5318899-26	Sequence 2, Appli
40	9	1.2	96	1	US-07-623-611-2	Sequence 3, Appli
41	9	1.2	96	1	US-07-623-611-3	Sequence 2, Appli
42	9	1.2	96	5	PCT-US91-09108-2	Sequence 3, Appli
43	9	1.2	96	5	PCT-US91-09108-3	Sequence 6, Appli
44	9	1.2	98	1	US-07-623-611-6	Sequence 9, Appli
45	9	1.2	98	1	US-07-623-611-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-08-765-243-8

; Sequence 8, Application US/08765243

; Patent No. 5935578

; GENERAL INFORMATION:

; APPLICANT: ALVES, KENNETH

; APPLICANT: GUPTA, SUNIL K.

; APPLICANT: HOLLIS, GREGORY F.

; TITLE OF INVENTION: CONTRACEPTIVE VACCINE

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MARY A. APPOLLINA

; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE

; CITY: RAHWAY

; STATE: NJ

; COUNTRY: USA

; ZIP: 07065

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/765,243

; FILING DATE:

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: APPOLLINA, MARY A

; REGISTRATION NUMBER: 34,087

; REFERENCE/DOCKET NUMBER: 19244Y

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (908)594-3462

; TELEFAX: (908)594-4720

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 734 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-765-243-8

Query Match	100.0%;	Score 734;	DB 2;	Length 734;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 734;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MWVFLSLGGLGRMDSNFDLPVQITVPKIRSIKEGIESQASYSKIVIEGPKPTVNLN	60	
Db	1	MWVFLSLGGLGRMDSNFDLPVQITVPKIRSIKEGIESQASYSKIVIEGPKPTVNLN	60	
Qy	61	QKNFLPHNFRVSYSGTGIMKPLDQDFNQFNCHYQGYIEGYPKSYVMVSTCTGLRGLQFE	120	

```
Db 1 QKNFLPHNFRVYSYSGTGIMKPLDQDFNQFYQYIEGYPKSVVMVSTCTGLRGLQFE 120
Qy 121 NVSYGIEPLESSVGFHVYIYQVHKKADVSLYNEKDIERSDLSPKLOSAPQODPAKYIE 180
Db 121 NVSYGIEPLESSVGFHVYIYQVHKKADVSLYNEKDIERSDLSPKLOSAPQODPAKYIE 180
Qy 181 MHVIVEKQLYNHMGSDTTVAQKVFQGLIGLITNAIFVSNITIIILSSLELWIDENKIATTG 240
Db 181 MHVIVEKQLYNHMGSDTTVAQKVFQGLIGLITNAIFVSNITIIILSSLELWIDENKIATTG 240
Qy 241 EANELHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMCDANYAGGVVLPHT 300
Db 241 EANELHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMCDANYAGGVVLPHT 300
Qy 301 ISLESVAVILAQLLSLNGITYDDINKCQCGVAVIMNPEAIHFSGVKIFSNCSFEDFAH 360
Db 301 ISLESVAVILAQLLSLNGITYDDINKCQCGVAVIMNPEAIHFSGVKIFSNCSFEDFAH 360
Qy 361 FISKQKSOCLHNPRLDPPFKQAVCGNAKLEAGECDCTEQCALIGETCCDIATCRF 420
Db 361 FISKQKSOCLHNPRLDPPFKQAVCGNAKLEAGECDCTEQCALIGETCCDIATCRF 420
Qy 421 KAGSNCAEGPCENCLFMSKERMCRPSFECDLPEYCNCGSSASCENHYVOTGHPGGLNQ 480
Db 421 KAGSNCAEGPCENCLFMSKERMCRPSFECDLPEYCNCGSSASCENHYVOTGHPGGLNQ 480
Qy 481 WICIDGVCMSGDQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540
Db 481 WICIDGVCMSGDQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540
Qy 541 CGKLI CKYVGFLLQIPRATIIYANISGHLCIAVEFASDHADSKQMWIKDGTSCGSKVC 600
Db 541 CGKLI CKYVGFLLQIPRATIIYANISGHLCIAVEFASDHADSKQMWIKDGTSCGSKVC 600
Qy 601 RNQRCVSSSYLYGDCITDKNDRGVGNKKHCHCSASYLPDSCVQSDLPWGGSIDSGNF 660
Db 601 RNQRCVSSSYLYGDCITDKNDRGVGNKKHCHCSASYLPDSCVQSDLPWGGSIDSGNF 660
Qy 661 PPVAIPARLPERRYENIYHSKPMRWPFLLFIPFFIIFCVLIAMVKNFORKKWRTEDY 720
Db 661 PPVAIPARLPERRYENIYHSKPMRWPFLLFIPFFIIFCVLIAMVKNFORKKWRTEDY 720
Qy 721 SSDEQPESESEPKG 734
Db 721 SSDEQPESESEPKG 734
```

## RESULT 2

```
PCT-US95-07295-8
; Sequence 8, Application PC/TUS9507295
; GENERAL INFORMATION:
; APPLICANT: ALVES, KENNETH
; APPLICANT: GUPTA, SUNIL K.
; APPLICANT: HOLLIS, GREGORY F.
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MARY A. APPOLLINA
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAYWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07295
; FILING DATE:
; CLASSIFICATION:
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: APPOLLINA, MARY A.
; REGISTRATION NUMBER: 34, 087
; REFERENCE/DOCKET NUMBER: 19244Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 734 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-07295-8
```

Query Match 100.0%; Score 734; DB 5; Length 734;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MWVLFLLSGLGLRMDNSFDSLPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNLM 60
Db 1 MWVLFLLSGLGLRMDNSFDSLPVQITVPEKIRSIIEGIESQASYKIVIEGKPYTVNLM 60
Qy 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFNQFYQYIEGYPKSVVMVSTCTGLRGLQFE 120
Db 61 QKNFLPHNFRVYSYSGTGIMKPLDQDFNQFYQYIEGYPKSVVMVSTCTGLRGLQFE 120
Qy 121 NVSYGIEPLESSVGFHVYIYQVHKKADVSLYNEKDIERSDLSPKLOSAPQODPAKYIE 180
Db 121 NVSYGIEPLESSVGFHVYIYQVHKKADVSLYNEKDIERSDLSPKLOSAPQODPAKYIE 180
Qy 181 MHVIVEKQLYNHMGSDTTVAQKVFQGLIGLITNAIFVSNITIIILSSLELWIDENKIATTG 240
Db 181 MHVIVEKQLYNHMGSDTTVAQKVFQGLIGLITNAIFVSNITIIILSSLELWIDENKIATTG 240
Qy 241 EANELHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMCDANYAGGVVLPHT 300
Db 241 EANELHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMCDANYAGGVVLPHT 300
Qy 301 ISLESVAVILAQLLSLNGITYDDINKCQCGVAVIMNPEAIHFSGVKIFSNCSFEDFAH 360
Db 301 ISLESVAVILAQLLSLNGITYDDINKCQCGVAVIMNPEAIHFSGVKIFSNCSFEDFAH 360
Qy 361 FISKQKSOCLHNPRLDPPFKQAVCGNAKLEAGECDCTEQCALIGETCCDIATCRF 420
Db 361 FISKQKSOCLHNPRLDPPFKQAVCGNAKLEAGECDCTEQCALIGETCCDIATCRF 420
Qy 421 KAGSNCAEGPCENCLFMSKERMCRPSFECDLPEYCNCGSSASCENHYVOTGHPGGLNQ 480
Db 421 KAGSNCAEGPCENCLFMSKERMCRPSFECDLPEYCNCGSSASCENHYVOTGHPGGLNQ 480
Qy 481 WICIDGVCMSGDQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540
Db 481 WICIDGVCMSGDQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540
Qy 541 CGKLI CKYVGFLLQIPRATIIYANISGHLCIAVEFASDHADSKQMWIKDGTSCGSKVC 600
Db 541 CGKLI CKYVGFLLQIPRATIIYANISGHLCIAVEFASDHADSKQMWIKDGTSCGSKVC 600
Qy 601 RNQRCVSSSYLYGDCITDKNDRGVGNKKHCHCSASYLPDSCVQSDLPWGGSIDSGNF 660
Db 601 RNQRCVSSSYLYGDCITDKNDRGVGNKKHCHCSASYLPDSCVQSDLPWGGSIDSGNF 660
Qy 661 PPVAIPARLPERRYENIYHSKPMRWPFLLFIPFFIIFCVLIAMVKNFORKKWRTEDY 720
Db 661 PPVAIPARLPERRYENIYHSKPMRWPFLLFIPFFIIFCVLIAMVKNFORKKWRTEDY 720
Qy 721 SSDEQPESESEPKG 734
Db 721 SSDEQPESESEPKG 734
```

RESULT 3

```
US-08-264-101-2
; Sequence 2, Application US/08264101
; Patent No. 5693496
; GENERAL INFORMATION:
; APPLICANT: ALVES, KENNETH
; APPLICANT: GUPTA, SUNIL K.
; APPLICANT: HOLLIS, GREGORY F.
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MARY A. APPOLLINA
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAYWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,101
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: APPOLLINA, MARY A.
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-264-101-2

Query Match      88.6%; Score 650; DB 1; Length 651;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      85 QDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVIYQVKH 144
DB      2 QDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVIYQVKH 61
QY      145 KKADVSLYNEKDIESRDLSPKLSAEPQDFAKYIEMHVIVEKQLYNHMGSDTTTVAQKV 204
DB      62 KKADVSLYNEKDIESRDLSPKLSAEPQDFAKYIEMHVIVEKQLYNHMGSDTTTVAQKV 121
QY      205 FQILGLTNAIFVSNITIISSLELWIDENKIATGSEANELLHTFLRWKTSYLVLRPHDV 264
DB      122 FQILGLTNAIFVSNITIISSLELWIDENKIATGSEANELLHTFLRWKTSYLVLRPHDV 181
QY      265 AFLVVRKSNYVGCATFQGWKQDANYAGGVVLPRTISLESIAVILAQLLSLNGITYDD 324
DB      182 AFLVVRKSNYVGCATFQGWKQDANYAGGVVLPRTISLESIAVILAQLLSLNGITYDD 241
QY      325 INKQCQSGAVICINPEAIHSGVKIFSNCSFEDFAHFIISKQSOCLHNQRLDPFFKQQA 384
DB      242 INKQCQSGAVICINPEAIHSGVKIFSNCSFEDFAHFIISKQSOCLHNQRLDPFFKQQA 301
QY      385 VCGNAKLEAGEBCDCGTEQDCAIGETCCDIATCRFKAGSNCAEGPCCNCLFMSKERM 444
DB      302 VCGNAKLEAGEBCDCGTEQDCAIGETCCDIATCRFKAGSNCAEGPCCNCLFMSKERM 361
QY      445 RPSFECDLPEYCNSSASCPENHYVQTGHPGCLNQWICIDGVCMSGDQCTDTFGKEVE 504
DB      362 RPSFECDLPEYCNSSASCPENHYVQTGHPGCLNQWICIDGVCMSGDQCTDTFGKEVE 421

US-08-765-243-2
; Sequence 2, Application US/08765243
; Patent No. 5935578
; GENERAL INFORMATION:
; APPLICANT: ALVES, KENNETH
; APPLICANT: GUPTA, SUNIL K.
; APPLICANT: HOLLIS, GREGORY F.
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MARY A. APPOLLINA
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAYWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,243
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: APPOLLINA, MARY A.
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19244Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-4720
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-765-243-2

Query Match      88.6%; Score 650; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      85 QDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVIYQVKH 144
DB      2 QDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVIYQVKH 61
QY      145 KKADVSLYNEKDIESRDLSPKLSAEPQDFAKYIEMHVIVEKQLYNHMGSDTTTVAQKV 204
DB      62 KKADVSLYNEKDIESRDLSPKLSAEPQDFAKYIEMHVIVEKQLYNHMGSDTTTVAQKV 121
QY      205 FQILGLTNAIFVSNITIISSLELWIDENKIATGSEANELLHTFLRWKTSYLVLRPHDV 264
DB      122 FQILGLTNAIFVSNITIISSLELWIDENKIATGSEANELLHTFLRWKTSYLVLRPHDV 181
QY      265 AFLVVRKSNYVGCATFQGWKQDANYAGGVVLPRTISLESIAVILAQLLSLNGITYDD 324
DB      182 AFLVVRKSNYVGCATFQGWKQDANYAGGVVLPRTISLESIAVILAQLLSLNGITYDD 241
QY      325 INKQCQSGAVICINPEAIHSGVKIFSNCSFEDFAHFIISKQSOCLHNQRLDPFFKQQA 384
DB      242 INKQCQSGAVICINPEAIHSGVKIFSNCSFEDFAHFIISKQSOCLHNQRLDPFFKQQA 301
QY      385 VCGNAKLEAGEBCDCGTEQDCAIGETCCDIATCRFKAGSNCAEGPCCNCLFMSKERM 444
DB      302 VCGNAKLEAGEBCDCGTEQDCAIGETCCDIATCRFKAGSNCAEGPCCNCLFMSKERM 361
QY      445 RPSFECDLPEYCNSSASCPENHYVQTGHPGCLNQWICIDGVCMSGDQCTDTFGKEVE 504
DB      362 RPSFECDLPEYCNSSASCPENHYVQTGHPGCLNQWICIDGVCMSGDQCTDTFGKEVE 421
```

Db 122 FQIIGLTNAIFVSNITIISSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 181  
QY 265 AFLVYREKSNYVGATFQGMKCDANYAGGVVHLHPTISLESVLAQLLSLNGIYDD 324  
Db 182 AFLVYREKSNYVGATFQGMKCDANYAGGVVHLHPTISLESVLAQLLSLNGIYDD 241  
QY 325 INKQCSCGAVCIMNPEAIHFSGVKIFSNCSFEDFAHFIKSKSOCLHNQRLDPPFFKQA 384  
Db 242 INKQCSCGAVCIMNPEAIHFSGVKIFSNCSFEDFAHFIKSKSOCLHNQRLDPPFFKQA 301  
QY 385 VCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERM 444  
Db 302 VCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERM 361  
QY 445 RPSFECDLPEYCNCGSSASCPENHYVOTGHPGCLNQCICIDGVCMSGDKCTDTFGKEVE 504  
Db 362 RPSFECDLPEYCNCGSSASCPENHYVOTGHPGCLNQCICIDGVCMSGDKCTDTFGKEVE 421  
QY 505 FGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLIKCYVGFLLQIPRATIIYA 564  
Db 422 FGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLIKCYVGFLLQIPRATIIYA 481  
QY 565 NISGHLCIAVEFASDHADSQMWIKDGTSCGSKNVCNQRCSVSSYLGVDCTTDKCNDRG 624  
Db 482 NISGHLCIAVEFASDHADSQMWIKDGTSCGSKNVCNQRCSVSSYLGVDCTTDKCNDRG 541  
QY 625 VCNNKXKHCASLYLPDCSVQSDLPWPGGSDSGNFPFPAIPARLPERRRYENIYHSKPM 684  
Db 542 VCNNKXKHCASLYLPDCSVQSDLPWPGGSDSGNFPFPAIPARLPERRRYENIYHSKPM 601  
QY 685 RWPFFLFIPIFCVLIAIMVKNFQKWKRTEDYSSDEQSESEPKG 734  
Db 602 RWPFFLFIPIFCVLIAIMVKNFQKWKRTEDYSSDEQSESEPKG 651

## RESULT 5

PCT-US95-07295-2  
; Sequence 2, Application PC/TUS9507295  
; GENERAL INFORMATION:  
; APPLICANT: ALVES, KENNETH  
; APPLICANT: GUPTA, SUNIL K.  
; APPLICANT: HOLLISS, GREGORY F.  
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MARY A. APOLLINA  
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE  
; CITY: RAHWAY  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/07295  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: APOLLINA, MARY A  
; REGISTRATION NUMBER: 34,087  
; REFERENCE/DOCKET NUMBER: 19244Y  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908)594-3462  
; TELEFAX: (908)594-4720  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 651 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
PCT-US95-07295-2  
Query Match 88.6%; Score 650; DB 5; Length 651;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 650; Conservative 0; Mismatches 0; Indels- 0; Gaps 0;  
QY 85 QDFQNFCHQGYIEGYPKSVVMYSTCTGLRGVLQFENVSYGIEPLLESSVGFHVIQVKH 144  
Db 2 QDFQNFCHQGYIEGYPKSVVMYSTCTGLRGVLQFENVSYGIEPLLESSVGFHVIQVKH 61  
QY 145 KKADVSLYNEKDIIESRDLSPKLSAPQDFAKYIEKHVIVEKQLXNHMGSDTTVAQKV 204  
Db 62 KKADVSLYNEKDIIESRDLSPKLSAPQDFAKYIEKHVIVEKQLXNHMGSDTTVAQKV 121  
QY 205 FQIIGLTNAIFVSNITIISSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 264  
Db 122 FQIIGLTNAIFVSNITIISSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 181  
QY 265 AFLVYREKSNYVGATFQGMKCDANYAGGVVHLHPTISLESVLAQLLSLNGIYDD 324  
Db 182 AFLVYREKSNYVGATFQGMKCDANYAGGVVHLHPTISLESVLAQLLSLNGIYDD 241  
QY 325 INKQCSCGAVCIMNPEAIHFSGVKIFSNCSFEDFAHFIKSKSOCLHNQRLDPPFFKQA 384  
Db 242 INKQCSCGAVCIMNPEAIHFSGVKIFSNCSFEDFAHFIKSKSOCLHNQRLDPPFFKQA 301  
QY 385 VCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERM 444  
Db 302 VCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERM 361  
QY 445 RPSFECDLPEYCNCGSSASCPENHYVOTGHPGCLNQCICIDGVCMSGDKCTDTFGKEVE 504  
Db 362 RPSFECDLPEYCNCGSSASCPENHYVOTGHPGCLNQCICIDGVCMSGDKCTDTFGKEVE 421  
QY 505 FGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLIKCYVGFLLQIPRATIIYA 564  
Db 422 FGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLIKCYVGFLLQIPRATIIYA 481  
QY 565 NISGHLCIAVEFASDHADSQMWIKDGTSCGSKNVCNQRCSVSSYLGVDCTTDKCNDRG 624  
Db 482 NISGHLCIAVEFASDHADSQMWIKDGTSCGSKNVCNQRCSVSSYLGVDCTTDKCNDRG 541  
QY 625 VCNNKXKHCASLYLPDCSVQSDLPWPGGSDSGNFPFPAIPARLPERRRYENIYHSKPM 684  
Db 542 VCNNKXKHCASLYLPDCSVQSDLPWPGGSDSGNFPFPAIPARLPERRRYENIYHSKPM 601  
QY 685 RWPFFLFIPIFCVLIAIMVKNFQKWKRTEDYSSDEQSESEPKG 734  
Db 602 RWPFFLFIPIFCVLIAIMVKNFQKWKRTEDYSSDEQSESEPKG 651

## RESULT 6

US-10-000-489-70  
; Sequence 70, Application US/10000489  
; Patent No. 6794363  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US6.DIV  
; CURRENT APPLICATION NUMBER: US/10/000,489  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: Jpatent  
; SEQ ID NO 70  
; LENGTH: 787  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: 1..16  
US-10-000-489-70

Query Match 2.0%; Score 15; DB 4; Length 787;  
Best Local Similarity 100.0%; Pred. No. 2.7e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 LQENVSYGIEPLES 131  
Db 121 LQENVSYGIEPLES 135

RESULT 7  
PCT-US93-05640-34  
; Sequence 34, Application PC/TUS9305640  
; GENERAL INFORMATION:  
; APPLICANT: University of Connecticut  
; TITLE OF INVENTION: Contraceptive Vaccine  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/05640  
; FILING DATE: 19930610  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/897,883  
; FILING DATE: June 12, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brook, David E.  
; REGISTRATION NUMBER: 22,592  
; REFERENCE/DOCKET NUMBER: UCT90-01AA  
; TELEPHONE: (617) 861-9540  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 48 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US93-05640-34

Query Match 1.6%; Score 12; DB 5; Length 48;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 605 CVSSSYLGIDYCT 616  
Db 33 CVSSSYLGIDYCT 44

RESULT 8  
PCT-US93-05640-16  
; Sequence 16, Application PC/TUS9305640  
; GENERAL INFORMATION:  
; APPLICANT: University of Connecticut  
; TITLE OF INVENTION: Contraceptive Vaccine  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/05640  
; FILING DATE: 19930610  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/897,883  
; FILING DATE: June 12, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brook, David E.  
; REGISTRATION NUMBER: 22,592  
; REFERENCE/DOCKET NUMBER: UCT90-01AA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-9540  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 50 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US93-05640-16

Query Match 1.8%; Score 12; DB 5; Length 50;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 450 ECDLPEYCNCGSS 461  
Db 26 ECDLPEYCNCGSS 37

RESULT 9  
PCT-US93-05640-25  
; Sequence 25, Application PC/TUS9305640  
; GENERAL INFORMATION:  
; APPLICANT: University of Connecticut  
; TITLE OF INVENTION: Contraceptive Vaccine  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/05640  
; FILING DATE: 19930610

```

QY      121 NVSYGIEPLESS 132
      |||||
Db      128 NVSYGIEPLESS 139

RESULT 11
PCT-US95-07295-6
; Sequence 6, Application PC/TUS9507295
; GENERAL INFORMATION:
; APPLICANT: ALVES, KENNETH
; APPLICANT: GUPTA, SUNIL K.
; APPLICANT: HOLLIS, GREGORY F.
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MARY A. APPOLLINA
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07295
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: APPOLLINA, MARY A.
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19244Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-07295-6

Query Match      1.6%; Score 12; DB 5; Length 735;
Best Local Similarity 100.0%; Pred.No. 0.0029;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      121 NVSYGIEPLESS 132
      |||||
Db      128 NVSYGIEPLESS 139

RESULT 12
PCT-US93-05640-12
; Sequence 12, Application PC/TUS9305640
; GENERAL INFORMATION:
; APPLICANT: University of Connecticut
; TITLE OF INVENTION: Contraceptive Vaccine
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS

```



```
;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05640
; FILING DATE: 19930610
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/897,883
; FILING DATE: June 12, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: UCT90-01AA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-05640-12

Query Match 1.4%; Score 10; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 LEAGECDCG 400
Db 10 LEAGECDCG 19

RESULT 13
PCT-US93-05640-23
; GENERAL INFORMATION:
; APPLICANT: University of Connecticut
; TITLE OF INVENTION: Contraceptive Vaccine
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05640
; FILING DATE: 19930610
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/897,883
; FILING DATE: June 12, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: UCT90-01AA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```

```
PCT-US93-05640-23

Query Match 1.4%; Score 10; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 LEAGECDCG 400
Db 10 LEAGECDCG 19

RESULT 14
US-08-993-165-16
; Sequence 16, Application US/08993165A
; Patent No. 6123923
; GENERAL INFORMATION:
; APPLICANT: Unger, Evan C
; APPLICANT: Wu, Yunqiu
; TITLE OF INVENTION: Optacoustic Contrast Agents And Methods For Their Use
; FILE REFERENCE: UNGR1224
; CURRENT APPLICATION NUMBER: US/08/993,165A
; CURRENT FILING DATE: 1997-12-18
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: No. 6123923el Sequence
US-08-993-165-16

Query Match 1.4%; Score 10; DB 3; Length 68;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 EAGECDCGT 401
Db 1 EAGECDCGT 10

RESULT 15
US-09-540-448-16
; Sequence 16, Application US/09540448
; Patent No. 6403056
; GENERAL INFORMATION:
; APPLICANT: Unger, Evan C.
; TITLE OF INVENTION: Charged Lipids and Uses For The Same
; FILE REFERENCE: UNGR1592
; CURRENT APPLICATION NUMBER: US/09/540,448
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 08/925,353
; PRIOR FILING DATE: 1997-09-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: No. 6403056el Sequence
US-09-540-448-16

Query Match 1.4%; Score 10; DB 4; Length 68;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 EAGECDCGT 401
Db 1 EAGECDCGT 10
```

Search completed: January 10, 2005, 22:39:42

Job time : 42 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:28:38 ; Search time 44 Seconds  
(without alignments)  
1605.071 Million cell updates/sec

Title: US-10-054-683-19  
Perfect score: 734  
Sequence: 1 MWVFLSLGLGRLMDSNF.....WRTEDYSDEQPESEPKG 734  
Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :  
1: PIR:79:.\*  
2: PIR:.\*  
3: PIR3:.\*  
4: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	734	100.0	734	2	JC4861
2	52	7.1	735	2	G02937
3	12	1.6	357	2	S23403
4	10	1.4	71	2	S13168
5	10	1.4	73	2	A43020
6	10	1.4	73	2	A43019
7	10	1.4	416	2	A37877
8	10	1.4	481	2	JC4342
9	10	1.4	481	2	S43125
10	10	1.4	736	2	S47645
11	9	1.2	71	2	A59412
12	9	1.2	71	2	A33990
13	9	1.2	72	2	F43019
14	9	1.2	72	2	B43020
15	9	1.2	73	2	A59410
16	9	1.2	73	2	C43019
17	9	1.2	73	2	E43019
18	9	1.2	73	2	B40003
19	9	1.2	73	2	A40003
20	9	1.2	524	2	S38539
21	9	1.2	655	2	JC7850
22	9	1.2	660	2	S71949
23	9	1.2	670	2	I65967
24	9	1.2	671	2	AE0486
25	9	1.2	823	2	S18968
26	8	1.1	71	2	G43019
27	8	1.1	72	2	B43019
28	8	1.1	72	2	D43019
29	8	1.1	72	2	I43019

ALIGNMENTS

RESULT 1

JC4861  
fertilin beta chain - human  
C;Species: Homo sapiens (man)  
C;Date: 15-Aug-1996 #sequence\_revision 18-Oct-1996 #text\_change 09-Jul-2004  
C;Accession: JC4861  
R;Gupta, S.K.; Alves, K.; O'Neill Palladino, L.; Mark, G.E.; Hollis, G.F.  
Biochem. Biophys. Res. Commun. 224, 318-326, 1996  
A;Title: Molecular cloning of the human fertilin beta subunit.  
A;Reference number: JC4861; MUID:96295488; PMID:8702389  
A;Accession: JC4861  
A;Molecule type: mRNA  
A;Residues: 1-734 <GUP>  
A;Cross-references: UNIPROT:Q99965; GB:U38805; NID:G4151118; PIDN:RAD04206.1; PID:G41511  
C;Comment: This protein is an integral sperm membrane glycoprotein, and plays a role in  
C;Superfamily: mouse meltrin alpha; disintegrin homology  
C;Keywords: glycoprotein; integrin binding; transmembrane protein  
F;382-734/Product: fertilin beta chain #status predicted <NAT>  
F;382-467/Domain: disintegrin homology <DIS>  
F;448-450/Region: integrin binding #status predicted  
F;686-708/Domain: transmembrane #status predicted <TM>  
F;121,219,352,458,565/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match		100.0%;	Score 734;	DB 2;	Length 734;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 734;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MWVFLSLGLGRLMDSNFD	SLPVQITVPKIRSI	KEGIESQASYKIVIEGKPYTNLM	60
Db	1	MWVFLSLGLGRLMDSNFD	SLPVQITVPKIRSI	KEGIESQASYKIVIEGKPYTNLM	60
Qy	61	QKNFLPHNFRVYSYSGTGMKPLDDQF	QNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFE	120	
Db	61	QKNFLPHNFRVYSYSGTGMKPLDDQF	QNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFE	120	
Qy	121	NVSYGIEPLESSYGFHVIYQVKHKADVSL	NYNEKDIESRDLSPFKLOSAPQDFAKYIE	180	
Db	121	NVSYGIEPLESSYGFHVIYQVKHKADVSL	NYNEKDIESRDLSPFKLOSAPQDFAKYIE	180	
Qy	181	MHVIVEKQLYNHMGSDTTVAQKVQLIGLT	NAIFVSFNITITLSSLELWIDENKIATTG	240	
Db	181	MHVIVEKQLYNHMGSDTTVAQKVQLIGLT	NAIFVSFNITITLSSLELWIDENKIATTG	240	
Qy	241	EANELHHTFLRWKTSYLVLRPHDVAFL	LYVREKSNYVGATFOGKMCDDANVAGGVVLRPT	300	
Db	241	EANELHHTFLRWKTSYLVLRPHDVAFL	LYVREKSNYVGATFOGKMCDDANVAGGVVLRPT	300	
Qy	301	ISLESIAVILAQLLSLMSGITDYDDINK	CCSGAVCIWNPFAIHFSGVKIFSNCSFEDFAH	360	
Db	301	ISLESIAVILAQLLSLMSGITDYDDINK	CCSGAVCIWNPFAIHFSGVKIFSNCSFEDFAH	360	
Qy	361	FISKQKSCQLHNQRLDPPFPKQAVCG	NAKLEAGEBCDCTEQDCALIGETCCDIATCRF	420	

Db 361 FISQKSQLHNPRLDPFFKQAVCGNAKLEAGEEDCGTEQCALITGETCCDIATCRF 420  
QY 421 KAGNCAEGPCCNCLFMSKERMCRPSEECDLPEYCNCGSSASCENHYVOTGHPCGLNQ 480  
Db 421 KAGNCAEGPCCNCLFMSKERMCRPSEECDLPEYCNCGSSASCENHYVOTGHPCGLNQ 480  
QY 481 WICIDGVCMSGDKQCTDTFGKEVFGPSECYSHLNSKTDVSGNGCIGSDSGYTCQEAADNLQ 540  
Db 481 WICIDGVCMSGDKQCTDTFGKEVFGPSECYSHLNSKTDVSGNGCIGSDSGYTCQEAADNLQ 540  
QY 541 CGKLICKYVGFLLQIPRATIIYINISGHLICIAVEFASDHADSKMKIKDGTSCGSKNVC 600  
Db 541 CGKLICKYVGFLLQIPRATIIYINISGHLICIAVEFASDHADSKMKIKDGTSCGSKNVC 600  
QY 601 RNORCVSSVLYGYDCTDKNDRCVGNCKKHCHCSASYLPDSCVQSDLPVGGSIDSGNF 660  
Db 601 RNORCVSSVLYGYDCTDKNDRCVGNCKKHCHCSASYLPDSCVQSDLPVGGSIDSGNF 660  
QY 661 PPVAIPARLPERRIENIYHSPMRWPFLLPFFIIFCVLIAMVKNVFORKKWRTEDY 720  
Db 661 PPVAIPARLPERRIENIYHSPMRWPFLLPFFIIFCVLIAMVKNVFORKKWRTEDY 720  
QY 721 SSDEQPSSESPKG 734  
Db 721 SSDEQPSSESPKG 734

## RESULT 2

G02937  
fertilin beta - crab-eating macaque  
C/Species: Macaca fascicularis (Crab-eating macaque)  
C/Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 09-Jul-2004  
C/Accession: G02937; S55061  
R/Ramarao, C.S.; Myles, D.G.; White, J.M.; Primakoff, P.  
submitted to the EMBL Data Library, August 1995  
A/Reference number: G12615  
A/Accession: G02937  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-735 <RAM>  
A/Cross-references: UNIPROT:Q28478; EMBL:U03959; NID:G998339; PID:G998340  
R/Perry, A.C.F.; Gichuhi, P.M.; Jones, R.; Hall, L.  
Biochem. J. 307, 843-850, 1995  
A/Title: Cloning and analysis of monkey fertilin reveals novel alpha subunit isoforms.  
A/Reference number: S55059; MUID:95260313; PMID:7741716  
A/Accession: S55061  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-722, 'S', 724-735 <PER>  
A/Cross-references: EMBL:X77653; NID:G794076; PIDN:CAA54733.1; PID:G794077  
C/Species: mouse meltrin alpha; disintegrin homology  
F/393-468/Domain: disintegrin homology <DIS>

Query Match 7.1%; Score 52; DB 2; Length 735;  
Best Local Similarity 100.0%; Pred. No. 7.6e-46;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFNVSYGIEPLESSVGFHVIYQVKKADVSLYNEKDIESRDLSPKLOS 168  
Db 118 LQFNVSYGIEPLESSVGFHVIYQVKKADVSLYNEKDIESRDLSPKLOS 169

## RESULT 3

S23403  
sperm surface protein PH-30 beta chain precursor - guinea pig (fragment)  
C/Species: Cavia porcellus (guinea pig)  
C/Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 09-Jun-2000  
C/Accession: S23403; S25696  
R/Blobel, C.P.; Wolfberg, T.G.; Turck, C.W.; Myles, D.G.; Primakoff, P.; White, J.M.  
Nature 356, 248-252, 1992  
A/Title: A potential fusion peptide and an integrin ligand domain in a protein active in  
A/Reference number: S23402; MUID:92204234; PMID:1552944

A/Accession: S23403  
A/Molecule type: mRNA  
A/Residues: 1-357 <BL01>  
A/Cross-references: EMBL:Z11720  
A/Accession: S25696  
A/Molecule type: protein  
A/Residues: 5-8, 'X', 10-32; 67-88; 125-129, 'X', 131-134, 'X', 136-141, 'X', 143; 154, 'X', 156-161;  
C/Species: mouse meltrin alpha; disintegrin homology  
C/Keywords: glycoprotein; transmembrane protein  
F/1-4/Domain: signal sequence (fragment) #status predicted <SIG>  
F/5-357/Product: sperm surface protein PH-30 beta chain #status experimental <MAT>  
F/5-87/Domain: disintegrin homology <DIS>  
F/303-323/Domain: transmembrane #status predicted <TMM>  
F/78,186,256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.6%; Score 12; DB 2; Length 357;  
Best Local Similarity 100.0%; Pred. No. 0.00055;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 450 ECDLPEYCNCGSS 461  
Db 70 ECDLPEYCNCGSS 81

## RESULT 4

S13168  
batroxostatin - barba amarilla  
C/Species: Bothrops atrox (barba amarilla, fer-de-lance)  
C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Aug-2004  
C/Accession: S13168  
R/Rucinski, B.; Niewiarowski, S.; Holt, J.C.; Soszka, T.; Knudsen, K.A.  
Biochim. Biophys. Acta 1054, 257-262, 1990  
A/Title: Batroxostatin, an Arg-Gly-Asp-containing peptide from Bothrops atrox, is a potent  
A/Reference number: S13168; MUID:91002685; PMID:2207176

A/Accession: S13168  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-71 <RUC>  
A/Cross-references: UNIPROT:P18618  
C/Species: disintegrin homology  
F/1-69/Domain: disintegrin homology (fragment) <DIS>  
F/51-53/Region: cell attachment (R-G-D) motif

Query Match 1.4%; Score 10; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 EAGECDCCGT 401  
Db 1 EAGECDCCGT 10

## RESULT 5

A43020  
platelet aggregation disintegrin (jararacin), venom - jararacussu  
C/Species: Bothrops jararacussu (jararacussu)  
C/Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 10-Jul-1998  
C/Accession: A43020  
R/Scarborough, R.M.; Rose, J.W.; Naughton, M.A.; Phillips, D.R.; Nannizzi, L.; Arfsten, J.  
J. Biol. Chem. 268, 1058-1065, 1993  
A/Title: Characterization of the integrin specificities of disintegrins isolated from Ame-  
A/Reference number: A43019; MUID:93123215; PMID:8419314  
A/Accession: A43020  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-73 <SCA>  
C/Species: unassigned disintegrins; disintegrin homology  
F/3-69/Domain: disintegrin homology <DIS>

Query Match 1.4%; Score 10; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 EAGEECDCGT 401  
Db 1 EAGEECDCGT 10

RESULT 6  
A43019  
platelet aggregation disintegrin (cerastin), venom - Mojave Desert sidewinder  
C;Species: Crotalus cerastes cerastes (Mojave Desert sidewinder)  
C;Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 09-Jul-2004  
C;Accession: A43019  
R;Scarborough, R.M.; Rose, J.W.; Naughton, M.A.; Phillips, L.; Arfsten, J. Biol. Chem. 268, 1058-1065, 1993  
A;Title: Characterization of the integrin specificities of disintegrins isolated from Am  
A;Reference number: A43019; MUID:93123215; PMID:8419314  
A;Accession: A43019  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-73 <SCA>  
A;Cross-references: UNIPROT:P31982  
C;Superfamily: unassigned disintegrins; disintegrin homology  
F;1-69/Domain: disintegrin homology (fragment) <Dis>

Query Match 1.4%; Score 10; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 EAGEECDCGT 401  
Db 1 EAGEECDCGT 10

RESULT 7  
A37877  
hemorrhagic protein HR1b - habu  
C;Species: Trimeresurus flavoviridis (habu)  
C;Date: 31-Jul-1991 #sequence\_revision 31-Jul-1991 #text\_change 09-Jul-2004  
C;Accession: A37877  
R;Takeya, H.; Oda, K.; Miyata, T.; Omori-Sato, T.; Iwanaga, S. J. Biol. Chem. 265, 16068-16073, 1990  
A;Title: The complete amino acid sequence of the high molecular mass hemorrhagic protein  
A;Reference number: A37877; MUID:90375459; PMID:2398046  
A;Accession: A37877  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-416 <TAK>  
C;Cross-references: UNIPROT:P20164  
C;Superfamily: mouse meltrin alpha: disintegrin homology  
F;210-292/Domain: disintegrin homology <Dis>  
F;144/Active site: Glu #status predicted

Query Match 1.4%; Score 10; DB 2; Length 416;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 LEAGEECDCG 400  
Db 219 LEAGEECDCG 228

RESULT 8  
JC4342  
fibrinolytic proteinase (EC 3.4.24.-) precursor - Chinese habu  
N;Alternate names: fibrinogenolytic proteinase  
C;Species: Trimeresurus mucrosquamatus (Chinese habu)  
C;Date: 21-Dec-1995 #sequence\_revision 08-Feb-1996 #text\_change 09-Jul-2004  
C;Accession: JC4342; PC4108  
R;Huang, K.F.; Hung, C.C.; Pan, F.M.; Chow, L.P.; Tsugita, A.; Chiou, S.H. Biochem. Biophys. Res. Commun. 216, 223-233, 1995  
A;Title: Characterization of multiple metalloproteinases with fibrinogenolytic activity analysis.  
A;Reference number: JC4342; MUID:96067555; PMID:7488093  
A;Accession: JC4342

A;Molecule type: mRNA  
A;Residues: 1-481 <HUA>  
A;Cross-references: UNIPROT:Q92119; EMBL:X91190  
A;Accession: PC4108  
A;Molecule type: protein  
A;Residues: 237-251 <HU2>  
A;Experimental source: venom gland  
C;Comment: The proteolytic activity is inhibited by EDTA or phenanthroline.  
C;Superfamily: trigramin precursor; disintegrin homology  
C;Keywords: anticoagulant; hydrolase; metalloproteinase; venom; zinc  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-187/Domain: propeptide #status predicted <PRO>  
F;188-392/Product: fibrinolytic proteinase #status predicted <MAT>  
F;399-477/Domain: disintegrin homology <Dis>  
F;411-478/Product: disintegrin #status predicted <DNT>  
F;459-461/Region: cell attachment (R-G-D) motif  
F;333,337,343/Binding site: zinc, catalytic (His) #status predicted  
F;334/Active site: Glu #status predicted

Query Match 1.4%; Score 10; DB 2; Length 481;  
Best Local Similarity 100.0%; Pred. No. 0.091;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 LEAGEECDCG 400  
Db 408 LEAGEECDCG 417

RESULT 9  
S43125  
trimucin precursor - Chinese habu  
N;Contains: hemorrhagic proteinase (EC 3.4.24.-); platelet aggregation inhibitor (disint  
C;Species: Trimeresurus mucrosquamatus (Chinese habu)  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S47570; S43125  
R;Tsai, I.H.; Wang, Y.M.; Lee, Y.H. Biochim. Biophys. Acta 1200, 337-340, 1994  
A;Title: Characterization of a cDNA encoding the precursor of platelet aggregation inhibi  
A;Reference number: S47570; MUID:94347824; PMID:8068721  
A;Accession: S47570  
A;Molecule type: mRNA  
A;Residues: 1-481 <TS2>  
A;Cross-references: UNIPROT:Q91505; EMBL:X77089; MID:9467703; PIDN:CAA54364.1; PID:94677  
C;Superfamily: trigramin precursor; disintegrin homology  
C;Keywords: anticoagulant; hydrolase; metalloproteinase; venom; zinc  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-187/Domain: propeptide #status predicted <PRO>  
F;188-392/Product: hemorrhagic proteinase #status predicted <HEM>  
F;399-477/Domain: disintegrin homology <Dis>  
F;411-478/Product: disintegrin #status predicted <DNT>  
F;459-461/Region: cell attachment (R-G-D) motif  
F;333,337,343/Binding site: zinc, catalytic (His) #status predicted  
F;334/Active site: Glu #status predicted

Query Match 1.4%; Score 10; DB 2; Length 481;  
Best Local Similarity 100.0%; Pred. No. 0.091;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 LEAGEECDCG 400  
Db 408 LEAGEECDCG 417

RESULT 10  
S47645  
MDC I protein - crab-eating macaque  
C;Species: Macaca fascicularis (crab-eating macaque)  
C;Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S47645  
R;Barker, H.L.; Perry, A.C.F.; Jones, R.; Hall, L. Biochim. Biophys. Acta 1218, 429-431, 1994  
A;Title: Sequence and expression of a monkey testicular transcript encoding TMDC I, a no  
A;Reference number: S47645; MUID:94325353; PMID:8049267

A;Accession: S47645  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-736 <BAR>  
A;Cross-references: UNIPROT:Q28482; EMBL:X76637; NID:G53S016; PIDN:CAA54085.1; PID:G5350  
C;Superfamily: mouse meltrin alpha; disintegrin homology  
F;392-477/Domain: disintegrin homology <DIS>

Query Match 1.4%; Score 10; DB 2; Length 736;  
Best Local Similarity 100.0%; Pred. No. 0.13; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

Qy 123 SYGIEPLESS 132  
| | | | | | | | | |  
Db 127 SYGIEPLESS 136

RESULT 11  
A59412  
KGD-bearing platelet-aggregation disintegrin (ussuristatin) 2 - Agkistrodon ussuriensis  
C;Species: Agkistrodon ussuriensis  
C;Date: 01-Mar-2002 #sequence\_revision 01-Mar-2002 #text\_change 09-Jul-2004  
C;Accession: A59412  
R;Terada, S.  
J. Biochem. 125, 31-35, 1999  
A;Title: Ussuristatin 2, A Novel KGD-Bearing Disintegrin from Agkistrodon ussuriensis Ve  
A;Reference number: A59412  
A;Accession: A59412  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-71 <TER>  
A;Cross-references: UNIPROT:Q7LZT4  
C;Keywords: anticoagulant; integrin inhibitor; venom

Query Match 1.2%; Score 9; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 0.2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 392 EAGEECDCG 400  
| | | | | | | | | |  
Db 1 EAGEECDCG 9

RESULT 12  
A33990  
aplaggin - eastern cottonmouth  
C;Species: Agkistrodon piscivorus piscivorus (eastern cottonmouth)  
C;Date: 23-Mar-1990 #sequence\_revision 23-Mar-1990 #text\_change 09-Jul-2004  
C;Accession: A33990  
R;Chao, B.H.; Jakubowski, J.A.; Savage, B.; Chow, E.P.; Marzec, U.M.; Harker, L.A.; Mara  
Proc. Natl. Acad. Sci. U.S.A. 86, 8050-8054, 1989  
A;Title: Agkistrodon piscivorus piscivorus platelet aggregation inhibitor: a potent inhi  
A;Reference number: A33990; MUID:90046735; PMID:2510158  
A;Accession: A33990  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-71 <CHA>  
A;Cross-references: UNIPROT:P16338  
C;Superfamily: unassigned disintegrins; disintegrin homology  
F;1-68/Domain: disintegrin homology (fragment) <DIS>

Query Match 1.2%; Score 9; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 0.2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 392 EAGEECDCG 400  
| | | | | | | | | |  
Db 1 EAGEECDCG 9

RESULT 13  
F43019  
platelet aggregation disintegrin (cotiarin), venom - cotiara

C;Species: Bothrops cotiara (cotiara)  
C;Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 16-Aug-2004  
C;Accession: F43019  
R;Scarborough, R.M.; Rose, J.W.; Naughton, M.A.; Phillips, D.R.; Nannizzi, L.; Arfsten, J.  
J. Biol. Chem. 268, 1058-1065, 1993  
A;Title: Characterization of the integrin specificities of disintegrins isolated from Ame  
C;Superfamily: mouse meltrin alpha; disintegrin homology  
A;Reference number: A43019; MUID:93123215; PMID:8419314  
A;Accession: F43019  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-72 <SCA>  
A;Cross-references: UNIPROT:P31988  
C;Superfamily: disintegrin homology  
F;1-69/Domain: disintegrin homology (fragment) <DIS>

Query Match 1.2%; Score 9; DB 2; Length 72;  
Best Local Similarity 100.0%; Pred. No. 0.2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 392 EAGEECDCG 400  
| | | | | | | | | |  
Db 1 EAGEECDCG 9

RESULT 14  
B43020  
platelet aggregation disintegrin (cereberin), venom - Arizona black rattlesnake  
C;Species: Crotalus viridis cerberus (Arizona black rattlesnake)  
C;Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 09-Jul-2004  
C;Accession: B43020  
R;Scarborough, R.M.; Rose, J.W.; Naughton, M.A.; Phillips, D.R.; Nannizzi, L.; Arfsten, J.  
J. Biol. Chem. 268, 1058-1065, 1993  
A;Title: Characterization of the integrin specificities of disintegrins isolated from Ame  
A;Reference number: A43019; MUID:93123215; PMID:8419314  
A;Accession: B43020  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-72 <SCA>  
A;Cross-references: UNIPROT:P31985  
C;Superfamily: unassigned disintegrins; disintegrin homology  
F;3-69/Domain: disintegrin homology <DIS>

Query Match 1.2%; Score 9; DB 2; Length 72;  
Best Local Similarity 100.0%; Pred. No. 0.2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 392 EAGEECDCG 400  
| | | | | | | | | |  
Db 1 EAGEECDCG 9

RESULT 15  
A59410  
platelet aggregation disintegrin (brevicaudin) 1b, venom - Gloydus halys brevicaudus  
C;Species: Gloydus halys brevicaudus  
C;Date: 01-Mar-2002 #sequence\_revision 01-Mar-2002 #text\_change 09-Jul-2004  
C;Accession: A59410  
R;Terada, S.  
Fukuoka Univ. Sci. Reports 30, 71-78, 2000  
A;Title: Isolation and Primary Structures of Platelet Aggregation Inhibitors from Gloydid  
A;Reference number: A59409  
A;Accession: A59410  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-73 <TER>  
A;Cross-references: UNIPROT:Q90220  
C;Keywords: anticoagulant; integrin inhibitor; venom

Query Match 1.2%; Score 9; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 0.2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 392 EAGEECDCG 400

Db           |||||||  
          1 EAGECDG 9

Search completed: January 10, 2005, 22:38:56  
Job time : 45 secs

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:27:53 ; Search time 202 Seconds  
(without alignments)  
2090.718 Million cell updates/sec

Title: US-10-054-683-19

Perfect score: 734

Sequence: 1 MWVFLSLGLGLRMDNSFD.....WRTEDYSSDEQPESESEPKG 734

Scoring table:

OLIGO Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	631	86.0	735	1 AD02 HUMAN	Q99865 homo sapien
2	241	32.8	579	2 Q6P2G0	Q6P290 homo sapien
3	241	32.8	579	2 AAH64547	AAH64547 homo sapi
4	52	7.1	735	1 AD02 MACFA	Q28478 macaca fasc
5	20	2.7	735	2 Q865F8	Q866a8 sus scrofa
6	20	2.7	731	1 AD02_RABIT	Q28660 oryctolagus
7	18	2.5	735	1 AD02_CAVPO	Q60411 cavia porce
8	16	2.2	746	1 AD18 MACFA	Q95194 macaca fasc
9	15	2.0	279	2 Q6ZWL1	Q6ZWL1 homo sapien
10	15	2.0	279	2 BAC85489	BAC85489 homo sapi
11	15	2.0	787	2 Q8TC27	Q8TC27 homo sapien
12	15	2.0	787	2 Q8TC42	Q8TC42 homo sapien
13	15	2.0	787	2 AAQ89099	AAQ89099 homo sapi
14	14	1.9	745	1 AD02_BOVIN	Q77780 bos taurus
15	13	1.8	677	2 Q8BJ80	Q8BJ80 mus musculus
16	13	1.8	719	1 AD18 MOUSE	Q9r157 mus musculus
17	13	1.8	730	2 Q80Y08	Q80Y08 mus musculus
18	13	1.8	737	1 AD02_RAT	Q83202 rattus norv
19	13	1.8	750	2 Q6P901	Q6P901 mus musculus
20	13	1.8	750	2 Q8K410	Q8K410 mus musculus
21	13	1.8	750	2 AAH60983	AAH60983 mus muscu
22	12	1.6	735	1 AD02 MOUSE	Q60718 mus musculus
23	12	1.6	873	2 Q42595	Q42595 xenopus lae
24	11	1.5	77	2 Q801Z4	Q801Z4 bothrops al
25	11	1.5	474	2 Q95LW7	Q95LW7 macaca fasc
26	10	1.4	71	1 DISI_BOTAT	P18618 bothrops at
27	10	1.4	73	1 DISI_BOTJA	P31989 bothrops ja
28	10	1.4	73	1 DISI_CROCC	P31982 crotalus ce
29	10	1.4	105	2 Q78CF2	Q78CF2 agkistrodon
30	10	1.4	317	2 Q90WC0	Q90WC0 agkistrodon
31	10	1.4	416	1 HR1B_TRIFL	P20164 trimereauru

#### RESULT 1

ID	AD02 HUMAN	STANDARD;	PRT;	735 AA.
DT	Q99955; P78326; Q9UQ08;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)			
DE	(Fertilin beta subunit) (PH-30) (PH30).			
GN	Names=ADAM2; Synonyms=FTNB;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Testis;			
RX	MEDLINE=97193554; PubMed=9041139;			
RA	Vidaeus C.M., von Kap-Herr C., Golden W.L., Eddy R.L., Shows T.B.,			
RA	Herr J.C.;			
RT	"Human fertilin beta: identification, characterization, and			
RT	chromosomal mapping of an ADAM gene family member.";			
RL	Mol. Reprod. Dev. 46:363-369(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Testis;			
RX	MEDLINE=96295488; PubMed=8702389;			
RA	Gupta S.K., Alves K., Palladino L.O., Mark G.E., Hollis G.F.;			
RT	"Molecular cloning of the human fertilin beta subunit.";			
RL	Biochem. Biophys. Res. Commun. 224:318-326(1996).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Testis;			
RX	MEDLINE=97224507; PubMed=9070941;			
RA	Burkin H.R., Burkin D.J., Davey P.M., Griffin D.K., Affara N.A.;			
RT	"Mapping, sequence, and expression analysis of the human fertilin beta			
RT	gene (FTNB).";			
RL	Genomics 40:190-192(1997).			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Testis;			
RA	Hall L., Frayne J.;			
RT	"Nucleotide sequence of the human fertilin beta transcript.";			
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Testis;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			

P14530 trimereauru  
Q90220 agkistrodon  
Q72269 trimereauru  
O57413 trimereauru  
Q90YA7 trimereauru  
Q91505 trimereauru  
Q90YA6 trimereauru  
P18619 trimereauru  
Q92119 trimereauru  
O73795 agkistrodon  
Q8J1R1 trimereauru  
Q7Q104 anopheles g  
Q28482 macaca fasc  
Q6TY20 mus musculus

#### ALIGNMENTS

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Kraywinetski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- FUNCTION: Sperm surface membrane protein that may be involved in  
CC sperm-egg plasma membrane adhesion and fusion during  
CC fertilization. Could have a direct role in sperm-zona binding or  
CC migration of sperm from the uterus into the oviduct. Interactions  
CC with egg membrane could be mediated via binding between its  
CC disintegrin-like domain to one or more integrins receptors on the  
CC egg. This is a non catalytic metalloprotease-like protein.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC -1- Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q99965-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q99965-2; Sequence=VSP\_005471;  
CC Note=No experimental confirmation available;  
CC -1- TISSUE SPECIFICITY: Expressed specifically in spermatogenic cells  
CC in the seminiferous cells. Not detected in fetal tissues.  
CC -1- DOMAIN: A tripeptide motif (FEE) within disintegrin-like domain  
CC could be involved in the binding to egg integrin receptor and thus  
CC could mediate sperm/egg binding.  
CC -1- PFM: The prodomain and the metalloprotease domain are cleaved  
CC during the epididymal maturation of the spermatozoa.  
CC -1- MISCELLANEOUS: In mammals, exists as a heterodimer composed of an  
CC alpha and beta subunits. In human, fertilin alpha is a pseudogene.  
CC -1- SIMILARITY: Belongs to peptidase family M12B.  
CC -1- SIMILARITY: Contains 1 disintegrin domain.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; U52370; AAC51110.1; -;  
CC EMBL; U38805; RAD04206.1; -;  
CC EMBL; X99374; CAAG7753.1; -;  
CC EMBL; AJ133005; CAB40813.1; -;  
CC EMBL; BC034957; AAH34957.1; -;  
CC PIR; JC4861; JC4861.  
CC HSP; P30403; IN4Y.  
CC MEROPS; M12\_950; -;  
CC Genew; HGNC:198; ADAM2.  
CC MIM; 601533; -;  
CC GO; GO:0005887; C: integral to plasma membrane; TAS.  
CC GO; GO:0005178; F: integrin binding; TAS.  
CC GO; GO:0008237; F: metalloprotease activity; TAS.  
CC GO; GO:0007342; P: fusion of sperm to egg plasma membrane; TAS.  
CC InterPro; IPR006586; ADAM\_cysteine.  
CC InterPro; IPR001762; Disintegrin.  
CC InterPro; IPR006209; EGF-like.  
CC InterPro; IPR001590; Peptidase M12B.  
CC InterPro; IPR002870; Peptidase M12B\_N.  
CC InterPro; IPR006025; Pept\_M\_2n\_BS.  
CC Pfam; PF00200; Disintegrin\_1.  
CC Pfam; PF01562; Pep\_M12B\_propep; 1.  
CC Pfam; PF01421; Reprolysin; 1.

DR PRINTS; PR00289; DISINTEGRIN.  
DR ProDom; PD000664; Disintegrin; 1.  
DR SWART; SM00608; ACR; 1.  
DR SMART; SM00050; DISIN; 1.  
DR PROSITE; PS00215; ADAM\_MEPRO; 1.  
DR PROSITE; PS00427; DISINTEGRIN\_1; 1.  
DR PROSITE; PS00214; DISINTEGRIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; FALSE NEG.  
DR PROSITE; PS01186; EGF\_2; FALSE NEG.  
KW Alternative splicing; Cell adhesion; EGF-like domain; Glycoprotein;  
KW Signal; Transmembrane.  
FT SIGNAL 1 16 Potential.  
FT PROPEP 17 174  
FT CHAIN 175 735 ADAM 2.  
FT DOMAIN 175 686 Extracellular (Potential).  
FT TRANSMEM 687 707 Potential.  
FT DOMAIN 708 735 Cytoplasmic (Potential).  
FT DOMAIN 175 382 Metalloprotease-like.  
FT DOMAIN 383 475 Disintegrin-like.  
FT DOMAIN 477 606 Cys-rich.  
FT DOMAIN 612 645 EGF-like.  
FT DISULFID 287 370 By similarity.  
FT DISULFID 329 354 By similarity.  
FT DISULFID 445 458 Potential.  
FT DISULFID 616 627 By similarity.  
FT DISULFID 621 633 By similarity.  
FT DISULFID 635 644 By similarity.  
FT CARBOHYD 122 122 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 220 220 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 353 353 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 459 459 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 566 566 N-linked (GlcNAc... ) (Potential).  
FT VARSPLIC 172 190 Missing (in isoform 2).  
FT FTId=VSP\_005471.  
FT CONFLICT 3 3 Missing (in Ref. 2).  
FT CONFLICT 21 21 D -> H (in Ref. 3).  
FT CONFLICT 99 99 E -> D (in Ref. 3).  
FT CONFLICT 106 106 V -> G (in Ref. 3).  
FT CONFLICT 170 170 V -> A (in Ref. 2).  
FT CONFLICT 288 288 D -> H (in Ref. 1).  
FT CONFLICT 321 321 I -> T (in Ref. 1).  
FT CONFLICT 388 388 G -> S (in Ref. 3).  
FT CONFLICT 396 398 EEC -> DEF (in Ref. 3).  
FT CONFLICT 501 501 G -> S (in Ref. 3).  
FT CONFLICT 529 529 D -> Y (in Ref. 3).  
FT CONFLICT 579 579 S -> G (in Ref. 3).  
FT CONFLICT 588 588 W -> L (in Ref. 3).  
FT CONFLICT 603 603 N -> D (in Ref. 3).  
FT CONFLICT 629 630 NK -> KQ (in Ref. 3).  
FT CONFLICT 638 638 S -> F (in Ref. 3).  
SQ SEQUENCE 735 AA; 82456 MW; 92867B5340BEE01F CRC64;  
Query Match 86.0%; Score 631; DB 1; Length 735;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 731; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 VLFSLGLGLRMDSNFDSLPVQITVPEKRSIIKEGIESQASQYKIVIEGKPYTNLMQK 62  
DB 4 VLFSLGLGLRMDSNFDSLPVQITVPEKRSIIKEGIESQASQYKIVIEGKPYTNLMQK 63  
QY 63 NPLPHFRVYSYSGTGIMKFLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLPENY 122  
DB 64 NPLPHFRVYSYSGTGIMKFLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLPENY 123  
QY 123 SYGIEPLESSVGFEHVIYQVKKHKADVSLYNEKDIESRDLSPFKLSAEPQDFAKYIEMH 182  
DB 124 SYGIEPLESSVGFEHVIYQVKKHKADVSLYNEKDIESRDLSPFKLSAEPQDFAKYIEMH 183  
QY 183 VIVEKQLYNHMGSDTTVAQVQLGLTNAIFVSNFNTIILSSLELWIDENKIATGEA 242  
DB 184 VIVEKQLYNHMGSDTTVAQVQLGLTNAIFVSNFNTIILSSLELWIDENKIATGEA 243  
QY 243 NELLHTFLRWKTSYLVLRPHDVAFLLVYREKSNVYGATFGKRCMDANYAGGVVLPRTIS 302

```
Db 244 NELLHTFLRWKTSVLRLPHDVAELLVREKSNVVGATFGQKMDANYAGGVVLPRTIS 303
Qy 303 LESLAVILAQLLSLSMGITYDDINKCQCSGAVCIMNPEAIHFGSKVIFNSCSFEDFAHFI 362
Db 304 LESLAVILAQLLSLSMGITYDDINKCQCSGAVCIMNPEAIHFGSKVIFNSCSFEDFAHFI 363
Qy 363 SKQSKQCLHNPRLDPFFKQAVCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRFKA 422
Db 364 SKQSKQCLHNPRLDPFFKQAVCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRFKA 423
Qy 423 GSNCAEGPCCNCLFMSKRCRPSFECDDLPEYCNSSASCPENHYVQTGHPCGLNQWI 482
Db 424 GSNCAEGPCCNCLFMSKRCRPSFECDDLPEYCNSSASCPENHYVQTGHPCGLNQWI 483
Qy 483 CIDGVCMSGDKQCTDTFTGKEVEFGPSECYSHLSKTDVSGNCGISDSGYTQCEADNLQCG 542
Db 484 CIDGVCMSGDKQCTDTFTGKEVEFGPSECYSHLSKTDVSGNCGISDSGYTQCEADNLQCG 543
Qy 543 KLIKCYVGKFLLOIPRATIIIVANISGHLCTAVEFASDHADSQKWKIKDGTSCGSKNVCRN 602
Db 544 KLIKCYVGKFLLOIPRATIIIVANISGHLCTAVEFASDHADSQKWKIKDGTSCGSKNVCRN 603
Qy 603 QRCVSSVLYGDCVTDKCNDRGVCKKCHCSASYLPPDCSVQSDLPWGGSIDSGNFP 662
Db 604 QRCVSSVLYGDCVTDKCNDRGVCKKCHCSASYLPPDCSVQSDLPWGGSIDSGNFP 663
Qy 663 VAIPARLPERRYENIYHSPMRPFFLFPFFIFVCLIAIMVKNVFORKKWRTEDYSS 722
Db 664 VAIPARLPERRYENIYHSPMRPFFLFPFFIFVCLIAIMVKNVFORKKWRTEDYSS 723
Qy 723 DEQPESESEPKG 734
Db 724 DEQPESESEPKG 735

RESULT 2
Q6P2G0 PRELIMINARY; PRT; 579 AA.
ID Q6P2G0
AC AAH64547
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE ADAM2 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RL Strausberg R.;
RL EMBL: BC064547; AAH64547.1; -.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR Pfam; PF00200; Disintegrin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reptolysin; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR PRODOM; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00500; DISIN; 1.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; 1.
DR PROSITE; PS00214; DISINTEGRIN_2; 1.
SQ SEQUENCE 579 AA; 64795 MW; 1F7AF0CDA2A064F9 CRC64;

Query Match 32.8%; Score 241; DB 2; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.7e-233; Mismatches 0; Indels 0; Gaps 0;
Matches 241; Conservative 0;

Qy 297 HPTISLESIAVILAQLLSLSMGITYDDINKCQCSGAVCIMNPEAIHFGSKVIFNSCSFE 356
Db 172 HPTISLESIAVILAQLLSLSMGITYDDINKCQCSGAVCIMNPEAIHFGSKVIFNSCSFE 231
Qy 357 DFAHFISKQSKQCLHNPRLDPFFKQAVCGNAKLEAGEBCDCGTEQDCALIGETCCDIA 416
Db 232 DFAHFISKQSKQCLHNPRLDPFFKQAVCGNAKLEAGEBCDCGTEQDCALIGETCCDIA 291
Qy 417 TCRFKAGSNCAEGPCCNCLFMSKRCRPSFECDDLPEYCNSSASCPENHYVQTGHPC 476
Db 292 TCRFKAGSNCAEGPCCNCLFMSKRCRPSFECDDLPEYCNSSASCPENHYVQTGHPC 351
Qy 477 GLNQWICIDGVCMSGDKQCTDTFTGKEVEFGPSECYSHLSKTDVSGNCGISDSGYTQCEA 536
Db 352 GLNQWICIDGVCMSGDKQCTDTFTGKEVEFGPSECYSHLSKTDVSGNCGISDSGYTQCEA 411
Qy 537 D 537
Db 412 D 412

RESULT 3
AAH64547 PRELIMINARY; PRT; 579 AA.
ID AAH64547
AC AAH64547;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE ADAM2 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
```

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallus D.B., Schmerch A., Schein J.B.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strauberg R.;  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC064547; A064547.1; -;  
SQ SEQUENCE 579 AA; 64795 MW; 177A90CDA2A064F99 CRC64;

Query Match 32.8%; Score 241; DB 2; Length 579;  
Best Local Similarity 100.0%; Pred. No. 1.7e-233;  
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 297 HPTISLESIAVLAQLLSMGITYDDINKKQCSGAVCINMPEAIHPSGKIFSNCSFE 356  
Db 172 HPTISLESIAVLAQLLSMGITYDDINKKQCSGAVCINMPEAIHPSGKIFSNCSFE 231  
Qy 357 DFAHFISKQSQCLHNPRLDPPFPKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIA 416  
Db 232 DFAHFISKQSQCLHNPRLDPPFPKQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIA 291  
Qy 417 TCRFKAGSNCAEGPCCNCLFMSKERMCRPSFECSDLPEYCNCGSSASCPENHYVQTGHPC 476  
Db 292 TCRFKAGSNCAEGPCCNCLFMSKERMCRPSFECSDLPEYCNCGSSASCPENHYVQTGHPC 351  
Qy 477 GLNQCICIDGVCMSGDKQCTDTFCKEVEFGSECVSHLNSKNTDVSNGCGISDSGYTQCEA 536  
Db 352 GLNQCICIDGVCMSGDKQCTDTFCKEVEFGSECVSHLNSKNTDVSNGCGISDSGYTQCEA 411  
Qy 537 D 537  
Db 412 D 412

RESULT 4  
AD02 MACFA STANDARD; PRT; 735 AA.  
AC Q28478: Q28472;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)  
DE (Fertilin beta subunit) (PH-30) (PH30) (PH30-beta).  
GN Name=ADAM2; Synonyms=FTNB;  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=95260313; PubMed=7741716;  
RA Perry A.C.F., Gichuhi P.M., Jones R., Hall L.;  
RT "Cloning and analysis of monkey fertilin reveals novel alpha subunit  
RT isoforms";  
RL Biochem. J. 307:843-850 (1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96341733; PubMed=8720115;  
RA Ramarao C.S., Myles D.G., White J.M., Primakoff P.;  
RT "Initial evaluation of fertilin as an immun contraceptive antigen and  
RT molecular cloning of the cynomolgus monkey fertilin beta subunit";  
RL Mol. Reprod. Dev. 43:70-75 (1996).  
CC -1- FUNCTION: Sperm surface membrane protein that may be involved in

CC sperm-egg plasma membrane adhesion and fusion during  
CC fertilization. Could have a direct role in sperm-zona binding or  
CC migration of sperm from the uterus into the oviduct. Interactions  
CC with egg membrane could be mediated via binding between its  
CC disintegrin-like domain to one or more integrins receptors on the  
CC egg. This is a non catalytic metalloprotease-like protein (By  
CC similarity).  
CC -1- SUBUNIT: Heterodimer with ADAM1/fertilin alpha.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: Expressed specifically in testis.  
CC -1- DOMAIN: A tripeptide motif (FDE) within disintegrin-like domain  
CC could be involved in the binding to egg integrin receptor and thus  
CC could mediate sperm/egg binding (By similarity).  
CC -1- PTM: The signal and the metalloprotease domain are cleaved during  
CC the epididymal maturation of the spermatozoa (By similarity).  
CC -1- SIMILARITY: Belongs to peptidase family M12B.  
CC -1- SIMILARITY: Contains 1 disintegrin domain.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X77653; CAA54733.1; -;  
CC EMBL: U33959; AAB52699.1; -;  
CC PIR: G02937; G02937.  
CC HSP: P30403; IN4Y.  
CC MEROPS: M12.950; -;  
CC InterPro: IPR006586; ADAM cysteine.  
CC InterPro: IPR001762; Disintegrin.  
CC InterPro: IPR000742; EGF\_2.  
CC InterPro: IPR006209; EGF\_like.  
CC InterPro: IPR001590; Peptidase\_M12B\_N.  
CC Pfam: PF00200; Disintegrin; 1.  
CC Pfam: PF01562; Pep M12B propep; 1.  
CC Pfam: PF01421; Reptolysin; 1.  
CC PRINTS: PR00289; DISINTEGRIN.  
CC PRODOM: PD000664; Disintegrin; 1.  
CC SMART: SM00608; ACR; 1.  
CC SMART: SM00050; DISIN; 1.  
CC PROSITE: PS50215; ADAM\_MBP; 1.  
CC PROSITE: PS00427; DISINTEGRIN\_1; 1.  
CC PROSITE: PS00214; DISINTEGRIN\_2; 1.  
CC PROSITE: PS00022; EGF\_1; FALSE\_NEG.  
CC PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
CC PROSITE: PS50026; EGF\_3; 1.  
CC Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
KW SIGNAL 1 16 Potential.  
FT PROPEP 17 174 By similarity.  
FT CHAIN 175 735 ADAM 2.  
FT DOMAIN 17 686 Extracellular (Potential).  
FT TRANSMEM 687 707 Potential.  
FT DOMAIN 708 735 Cytoplasmic (Potential).  
FT DOMAIN 175 382 Metalloprotease-like.  
FT DOMAIN 383 475 Disintegrin-like.  
FT DOMAIN 476 606 Cys-rich.  
FT DOMAIN 612 645 EGF-like.  
FT DISULFID 287 370 By similarity.  
FT DISULFID 329 354 By similarity.  
FT DISULFID 445 458 Potential.  
FT DISULFID 616 627 By similarity.  
FT DISULFID 621 633 By similarity.  
FT DISULFID 635 644 By similarity.  
FT CARBOHYD 76 76 N-linked (GLNAC. . .) (Potential).  
FT CARBOHYD 122 122 N-linked (GLNAC. . .) (Potential).  
FT CARBOHYD 220 220 N-linked (GLNAC. . .) (Potential).  
FT CARBOHYD 353 353 N-linked (GLNAC. . .) (Potential).  
FT CARBOHYD 459 459 N-linked (GLNAC. . .) (Potential).

```

FT CARBOHYD 566 566 N-linked (GlcNAc... ) (Potential).
FT CONFLICT 723 S -> T (in Ref. 2).
SQ SEQUENCE 735 AA; 82357 MW; BE84BE115E1D0B53 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 7.4e-43; Length 735;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQPNVSYGIEPLESSVGFEHVIYQVKHKADVSLYNEKDIESRDLSPFKLQS 168
|||||
DQ 118 LQPNVSYGIEPLESSVGFEHVIYQVKHKADVSLYNEKDIESRDLSPFKLQS 169
|||||

RESULT 5
Q866A8 PRELIMINARY; PRT; 735 AA.
ID AC Q866A8;
DC Q866A8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fertilin beta.
GN Name=FTNB;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22872989; PubMed=14510675;
RA Day A.E., Quilter C.R., Sargent C.A., Mileham A.J.;
RT "Chromosomal mapping, sequence and transcription analysis of the
RL porcine fertilin beta gene (ADAM2).";
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Day A.;
RL Thesis (2004), Department of Applied Sciences, Anglia Polytechnic
RL University, Cambridge, United Kingdom.
DR EMBL: AJ309003; CAC84225.2; -.
DR HSSP: P18619; 1FVL.
DR MEROPS: M12.950; -.
DR GO: GO:004222; F:metalloendopeptidase activity; IEA.
DR GO: GO:003676; F:nucleic acid binding; IEA.
DR GO: GO:006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR006586; ADAM cysteine.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR008994; Nucleic acid OB.
DR InterPro: IPR001590; Peptidase_M12B.
DR InterPro: IPR002870; Peptidase_M12B_N.
DR Pfam: PF00200; Disintegrin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR Pfam: PF01421; Repolysin; 1.
DR PRINTS: PS00289; Disintegrin.
DR ProDom: PD000664; Disintegrin; 1.
DR SMART: SM00608; ACR; 1.
DR SMART: SM00050; DISIN; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS0215; ADAM MEPRO; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; 1.
DR PROSITE: PS00214; DISINTEGRIN_2; 1.
DR PROSITE: PS0026; EGF_3; 1.
SQ SEQUENCE 735 AA; 81804 MW; 16F5554690AEF4AB CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.4e-10; Length 735;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 SFNITILSSLELWIDENKI 236
|||||

```

```

Db 218 SFNITILSSLELWIDENKI 237

RESULT 6
AD02 RABIT
ID AD02_RABIT STANDARD; PRT; 751 AA.
AC Q28660;
DC Q28660;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)
DE (Fertilin beta subunit) (PH-30) (PH30) (PH30-beta).
GN Name=ADAM2; Synonyms=FTNB;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97071141; PubMed=8914066;
RX Hardy C.M., Holland M.K.;
RT "Cloning and expression of recombinant rabbit fertilin.";
RL Mol. Reprod. Dev. 45:107-116(1996).
CC -!- FUNCTION: Sperm surface membrane protein that may be involved in
CC sperm-egg plasma membrane adhesion and fusion during
CC fertilization. Could have a direct role in sperm-zona binding or
CC migration of sperm from the uterus into the oviduct. Interactions
CC with egg membrane could be mediated via binding between its
CC disintegrin-like domain to one or more integrins receptors on the
CC egg. This is a non catalytic metalloprotease-like protein (By
CC similarity).
CC -!- SUBUNIT: Heterodimer with ADAM1/fertilin alpha.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed specifically in testis.
CC -!- DOMAIN: A tripeptide motif (VGE) within disintegrin-like domain
CC could be involved in the binding to egg integrin receptor and thus
CC could mediate sperm/egg binding (By similarity).
CC -!- FTM: The signal and the metalloprotease domain are cleaved during
CC the epididymal maturation of the spermatozoa (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U46070; AAA93321.1; -.
DR HSSP: P30403; 1NAY.
DR MEROPS: M12.950; -.
DR InterPro: IPR006586; ADAM_cysteine.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR008994; Nucleic_acid_OB.
DR InterPro: IPR001590; Peptidase_M12B.
DR InterPro: IPR002870; Peptidase_M12B_N.
DR Pfam: PF00200; Disintegrin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR Pfam: PF01421; Repolysin; 1.
DR PRINTS: PR00289; DISINTEGRIN.
DR ProDom: PD000664; Disintegrin; 1.
DR SMART: SM00608; ACR; 1.
DR SMART: SM00050; DISIN; 1.
DR PROSITE: PS0215; ADAM MEPRO; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; 1.
DR PROSITE: PS00214; DISINTEGRIN_2; 1.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; FALSE_NEG.

```

DR PROSITE; PS50026; EGF\_3; 1.  
KW Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
FT SIGNAL 1 16 Potential.  
FT PROPEP 17 173 By similarity.  
FT CHAIN 17 751 ADAM 2.  
FT DOMAIN 17 702 Extracellular (Potential).  
FT TRANSMEM 703 723 Potential.  
FT DOMAIN 724 751 Cytoplasmic (Potential).  
FT DOMAIN 174 381 Metalloprotease-like.  
FT DOMAIN 382 474 Disintegrin-like.  
FT DOMAIN 475 606 Cys-rich.  
FT DOMAIN 612 645 EGF-like.  
FT DISULFID 286 369 By similarity.  
FT DISULFID 328 353 By similarity.  
FT DISULFID 444 457 Potential.  
FT DISULFID 616 627 By similarity.  
FT DISULFID 621 633 By similarity.  
FT DISULFID 635 644 By similarity.  
FT CARBOHYD 122 147 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 147 147 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 219 219 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 289 289 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 352 352 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 434 434 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 458 458 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 559 559 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 566 566 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 643 643 N-linked (GlcNAc... ) (Potential).  
SQ SEQUENCE 751 AA; 83616 MW; 824C6FBFC4A9FCE1 CRC64;

Query Match 2.7%; Score 20; DB 1; Length 751;  
Best Local Similarity 100.0%; Pred. No. 1.4e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 SPNITILSSLEWIDENKI 236  
|||||  
Db 217 SPNITILSSLEWIDENKI 236

RESULT 7  
AD02\_CAVPO STANDARD; PRT; 735 AA.  
AC Q60411;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 03-JUN-2004 (Rel. 44, Last annotation update)  
DE ADAM 2 precursor (A disintegrin and metalloproteinase domain 2)  
DE (Fertilin beta subunit) (PH-30) (PH30).  
GN Name=ADAM2; Synonyms=FTNB;  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=94068486; PubMed=8248170;  
RA Wolfsberg T.G., Bazan J.F., Blobel C.P., Myles D.G., Primakoff P.,  
RA White J.M.;  
RA "The precursor region of a protein active in sperm-egg fusion contains  
RT a metalloprotease and a disintegrin domain: structural, functional,  
RT and evolutionary implications.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:10783-10787(1993).  
RN [2]  
RP SEQUENCE OF 383-735 FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=92204234; PubMed=1552944;  
RA Blobel C.P., Wolfsberg T.G., Turck C.W., Myles D.G., Primakoff P.,  
RA White J.M.;  
RA "A potential fusion peptide and an integrin ligand domain in a protein  
RT active in sperm-egg fusion.";  
RL Nature 356:248-252(1992).  
CC -1- FUNCTION: Sperm surface membrane protein that may be involved in

sperm-egg plasma membrane adhesion and fusion during fertilization. Could have a direct role in sperm-zona binding or migration of sperm from the uterus into the oviduct. Interactions with egg membrane could be mediated via binding between its disintegrin-like domain to one or more integrins receptors on the egg. This is a non catalytic metalloprotease-like protein (By similarity).  
CC SUBUNIT: Heterodimer with ADAM1/fertilin alpha.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: Expressed specifically in testis.  
CC -1- DEVELOPMENTAL STAGE: Expression begins during meiotic prophase.  
CC -1- DOMAIN: A tripeptide motif (TDE) within disintegrin-like domain could be involved in the binding to egg integrin receptor and thus could mediate sperm/egg binding (By similarity).  
CC -1- PTM: The signal and the metalloprotease domain are cleaved during the epididymal maturation of the spermatozoa (By similarity).  
CC -1- SIMILARITY: Belongs to peptidase family M12B.  
CC -1- SIMILARITY: Contains 1 disintegrin domain.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z11720; CAA77784.1; -.  
CC HSSP; P18619; 1FVL.  
DR MEROPS; M12.950; -.  
DR InterPro; IPR006586; ADAM cysteine.  
DR InterPro; IPR001762; Disintegrin.  
DR InterPro; IPR000742; EGF 2.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR001590; Peptidase M12B.  
DR InterPro; IPR002870; Peptidase M12B\_N.  
DR Pfam; PF02000; Disintegrin; 1.  
DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
DR Pfam; PF01421; Reprolysin; 1.  
DR PRINTS; PR00289; DISINTEGRIN.  
DR PRODOM; SM00664; Disintegrin; 1.  
DR SMART; SM00608; ACR; 1.  
DR SMART; SM00050; DISIN; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
DR PROSITE; PS00427; DISINTEGRIN\_1; 1.  
DR PROSITE; PS50214; DISINTEGRIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
DR PROSITE; PS50026; EGF\_3; 1.  
KW Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
FT SIGNAL 1 15 Potential.  
FT PROPEP 16 176 By similarity.  
FT CHAIN 177 735 ADAM 2.  
FT DOMAIN 16 680 Extracellular (Potential).  
FT TRANSMEM 681 701 Potential.  
FT DOMAIN 702 735 Cytoplasmic (Potential).  
FT DOMAIN 178 382 Metalloprotease-like.  
FT DOMAIN 383 472 Disintegrin-like.  
FT DOMAIN 473 609 Cys-rich.  
FT DOMAIN 610 643 EGF-like.  
FT DISULFID 287 370 By similarity.  
FT DISULFID 329 354 By similarity.  
FT DISULFID 442 455 Potential.  
FT DISULFID 614 625 By similarity.  
FT DISULFID 619 631 By similarity.  
FT DISULFID 633 642 By similarity.  
SQ SEQUENCE 735 AA; 81904 MW; 7535FC39F44FB645 CRC64;

Query Match 2.5%; Score 18; DB 1; Length 735;  
Best Local Similarity 100.0%; Pred. No. 1.4e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 SYLVLRPHDVAFLVYRE 272  
 |||||  
 Db 256 SYLVLRPHDVAFLVYRE 273

RESULT 8  
 AD18 MACFA STANDARD; PRT; 746 AA.  
 AC Q95194;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE ADAM 18 precursor (A disintegrin and metalloproteinase domain 18)  
 DE (Transmembrane metalloproteinase-like, disintegrin-like, and cysteine-rich protein III) (tmcd III).  
 GN Name=ADAM18; Synonyms=TMDC3;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=98328566; PubMed=9665629;  
 RA Rayne J., Jure J.A., Barker H.L., Perry A.C.F., Jones R., Hall L.;  
 RT "Macaque MDC family of proteins: sequence analysis, tissue distribution and processing in the male reproductive tract";  
 RL Mol. Hum. Reprod. 4:429-437(1998).  
 CC -1- FUNCTION: Sperm surface membrane protein that may be involved in spermatogenesis and fertilization. This is a non catalytic metalloproteinase-like protein (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Expressed predominantly in adult and prepubertal testis.  
 CC -1- DOMAIN: A tripeptide motif (ECD) within disintegrin-like domain could be involved in the binding to egg integrin receptor and thus could mediate sperm/egg binding (By similarity).  
 CC -1- PTM: The prodomain and the metalloproteinase-like domain are cleaved during the epididymal maturation of the spermatozoa.  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin domain.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; Y08617; CAA69909.1; -;  
 DR HSSP; P83469; IMPZ.  
 DR MEROPS; M12.957; -;  
 DR InterPro; IPR006586; ADAM cysteine.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR000742; EGF 2.  
 DR InterPro; IPR006209; EGF like.  
 DR InterPro; IPR001590; Peptidase M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR Pfam; PF00200; Disintegrin; 1.  
 DR Pfam; PF01562; Pep M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR ProDom; PD000664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR PROSITE; PS00215; ADAM\_MERPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN 1; FALSE\_NEG.  
 DR PROSITE; PS00214; DISINTEGRIN 2; 1.  
 DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.

DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PS0026; EGF\_3; 1.  
 KW EGF-like domain; Glycoprotein; Signal; Spermatogenesis; Transmembrane.  
 FT SIGNAL 1 16 Potential.  
 FT PROPEP 17 183 Potential.  
 FT CHAIN 184 746 ADAM 18.  
 FT DOMAIN 177 687 Extracellular (Potential).  
 FT TRANSMEM 688 708 Potential.  
 FT DOMAIN 709 746 Cytoplasmic (Potential).  
 FT DOMAIN 184 389 Metalloprotease-like.  
 FT DOMAIN 390 479 Disintegrin-like.  
 FT DOMAIN 480 619 Cys-rich.  
 FT DOMAIN 620 654 EGF-like.  
 FT DISULFID 293 376 By similarity.  
 FT DISULFID 335 360 By similarity.  
 FT DISULFID 458 483 Potential.  
 FT DISULFID 624 636 By similarity.  
 FT DISULFID 630 642 By similarity.  
 FT DISULFID 644 653 By similarity.  
 FT CARBOHYD 36 36 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 122 122 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 149 149 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 156 156 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 177 177 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 294 294 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 359 359 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 465 465 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 611 611 N-linked (GLNAC. .) (Potential).  
 FT CARBOHYD 625 625 N-linked (GLNAC. .) (Potential).  
 SQ SEQUENCE 746 AA; 84448 MW; 1D8C7E9071502E30 CRC64;  
 Query Match 2.2%; Score 16; DB 1; Length 746;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 117 LQFENVSYGIEPLESS 132  
 |||||  
 Db 118 LQFENVSYGIEPLESS 133

RESULT 9  
 Q6ZWLI PRELIMINARY; PRT; 279 AA.  
 ID Q6ZWLI;  
 AC Q6ZWLI;  
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)  
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)  
 DT 03-JUL-2004 (T-EMBLrel. 27, Last annotation update)  
 DE Hypothetical protein FLJ16017.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,  
 RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,  
 RA Kaeriyama Y., Moriya S., Chiba E., Momiya H., Onogawa S.,  
 RA Raga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,  
 RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,  
 RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,  
 RA Kimura K., Yamashita H., Mateuo K., Nakamura Y., Sekine M.,  
 RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,  
 RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,  
 RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK122608; BAC85489.1; -;  
 DR InterPro; IPR001590; Peptidase M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR Pfam; PF01562; Pep M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR PROSITE; PS00215; ADAM\_MERPRO; 1.





```
Query Match      2.0%; Score 15; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR PROSITE; PS50214; DISINTEGRIN 2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN 1.
DR PROSITE; PS00026; EGF_3; 1.
KW Integrin; Metalloprotease; Protease.
SQ SEQUENCE 787 AA; 87931 MW; CE0F54A50466B336 CRC64;

Query Match      2.0%; Score 15; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFENVSYGIEPLES 131
DB 121 LQFENVSYGIEPLES 135

RESULT 12
Q8TC42 PRELIMINARY; PRT; 787 AA.
AC Q8TC42;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE A disintegrin and metalloprotease domain 32.
GN Name=ADAM32;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Guaratine P.H.,
RA Richards S., Worley K.C., Hate S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinska M.I., Skalka U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026085; AAH26085.1; -.
DR HSSP; P30403; IN4Y.
DR MEROPS; M12.960; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR Pfam; PF00200; Disintegrin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
```

RC TISSUE=Testis;  
 RX MEDLINE=97304362; PubMed=9160725;  
 RA Waters S.I., White J.M.;  
 RT "Biochemical and molecular characterization of bovine fertilin alpha  
 and beta (ADAM 1 and ADAM 2): a candidate sperm-egg binding/fusion  
 complex.";  
 RL Biol. Reprod. 56:1245-1254(1997).  
 CC -!- FUNCTION: Sperm surface membrane protein that may be involved in  
 sperm-egg plasma membrane adhesion and fusion during  
 fertilization. Could have a direct role in sperm-zona binding or  
 migration of sperm from the uterus into the oviduct. Interactions  
 with egg membrane could be mediated via binding between its  
 disintegrin-like domain to one or more integrins receptors on the  
 egg. This is a non catalytic metalloprotease-like protein (By  
 similarity).  
 CC -!- SUBUNIT: Heterodimer with ADAM1/fertilin alpha.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Expressed specifically in testis.  
 CC -!- DOMAIN: A tripeptide motif (TDE) within disintegrin-like domain  
 could be involved in the binding to egg integrin receptor and thus  
 could mediate sperm/egg binding (By similarity).  
 CC -!- PTM: The signal and the metalloprotease domain are cleaved during  
 the epididymal maturation of the spermatozoa (By similarity).  
 CC -!- SIMILARITY: Belongs to peptidase family M12B.  
 CC -!- SIMILARITY: Contains 1 disintegrin domain.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR ENBL; AF086808; AAC62753.1; -  
 DR HSSP; P18619; 1PVL.  
 DR MEROPS; M12.950; -  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR008994; Nucleic acid OB.  
 DR InterPro; IPR001590; Peptidase M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR Pfam; PF00200; Disintegrin; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propesp; 1.  
 DR Pfam; PF01421; Reprolysein; 1.  
 DR PRINTS; PR00289; DISINTEGRIN.  
 DR ProDom; PD000664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS00215; ADAM\_MEROP; 1.  
 DR PROSITE; PS00427; DISINTEGRIN 1; 1.  
 DR PROSITE; PS00214; DISINTEGRIN 2; 1.  
 DR PROSITE; PS00022; EGF 1; FALSE NEG.  
 DR PROSITE; PS01186; EGF 2; FALSE NEG.  
 DR PROSITE; PS00025; EGF 3; 1.  
 KW Cell adhesion; EGF-like domain; Glycoprotein; Signal; Transmembrane.  
 FT SIGNAL 1 16  
 FT PROPEP 17 174 Potential.  
 FT CHAIN 175 745 By similarity.  
 FT DOMAIN 17 696 ADAM 2.  
 FT TRANSMEM 697 717 Extracellular (Potential).  
 FT DOMAIN 718 745 Potential.  
 FT DOMAIN 175 382 Cytoplasmic (Potential).  
 FT DOMAIN 383 474 Metalloprotease-like.  
 FT DOMAIN 475 605 Disintegrin-like.  
 FT DOMAIN 606 644 Cys-rich.  
 FT DOMAIN 611 644 EGF-like.  
 FT DOMAIN 148 151 Poly-Ser.  
 FT DISULFID 287 370 By similarity.

FT DISULFID 329 354 By similarity.  
 FT DISULFID 444 457 Potential.  
 FT DISULFID 615 626 By similarity.  
 FT DISULFID 620 632 By similarity.  
 FT DISULFID 634 643 By similarity.  
 FT CARBOHYD 122 122 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 220 220 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 353 353 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 458 458 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 558 558 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 745 AA; 83150 MW; B5D9DC0168999B00 CRC64;  
 Query Match 1.9%; Score 14; DB 1; Length 745;  
 Best Local Similarity 100.0%; Pred. No. 0.00015;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 117 LOFENVSYGIEPLE 130  
 Db 118 LOFENVSYGIEPLE 131  
 RESULT 15  
 Q8BJ80 PRELIMINARY; PRT; 677 AA.  
 ID Q8BJ80 AC Q8BJ80; 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE Mus musculus adult male testis cDNA, RIKEN full-length enriched  
 library, clone:170008E04 Product:METALLOPROTEASE/DISINTEGRIN  
 DE homolog.  
 GN Name=Adam32;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Testis;  
 RA The FANTOM Consortium,  
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Tashiro H., Itoh M.,

RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RT WIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer.,  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hangaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,  
RA Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK005759; BAC25124.1; -.  
DR HSSP; P18619; 1FVL.  
DR MGD; MGI:2653822; Adam32.  
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR006586; ADAM\_cysteine.  
DR InterPro; IPR001762; Disintegrin.  
DR InterPro; IPR001590; Peptidase\_M12B.  
DR InterPro; IPR002870; Peptidase\_M12B\_N.  
DR Pfam; PF00200; Disintegrin; 1.  
DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
DR Pfam; PF01421; Reprolysin; 1.  
DR ProDom; PD000664; Disintegrin; 1.  
DR SMART; SM00608; ACR; 1.  
DR SMART; SM00050; DISIN; 1.  
DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
DR PROSITE; PS50214; DISINTEGRIN\_2; 1.  
KW Integrin; Metalloprotease; Protease.  
SQ SEQUENCE 677 AA; 75934 MW; 985CD38F26EDCF9F CRC64;

Query Match 1.8%; Score 13; DB 2; Length 677;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 LQFENVSYGIEPL 129  
Db |||||  
122 LQFENVSYGIEPL 134

Search completed: January 10, 2005, 22:38:07  
Job time : 204 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:09:32 ; Search time 40 Seconds  
(without alignments)  
1216.935 Million cell updates/sec

Title: US-10-054-683-19

Perfect score: 3984

Sequence: 1 MWVFLSLGLGRLMDSNF.....WRTDYSSDEQPSSEPKG 734

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3984	100.0	734	2	US-08-765-243-8
2	3984	100.0	734	5	PCT-US95-07295-8
3	3551	89.1	651	1	US-08-264-101-2
4	3551	89.1	651	2	US-08-765-243-2
5	3551	89.1	651	5	PCT-US95-07295-2
6	2440	61.2	735	2	US-08-765-243-6
7	2440	61.2	735	5	PCT-US95-07295-6
8	1512	38.0	457	1	US-08-264-101-4
9	1512	38.0	457	2	US-08-765-243-4
10	1512	38.0	457	5	PCT-US95-07295-4
11	1506.5	37.8	787	4	US-10-000-489-70
12	1493.5	37.5	715	4	US-10-140-002-116
13	948.5	23.8	790	4	US-10-140-002-204
14	937	23.5	722	4	US-09-617-145-2
15	843.5	21.2	775	4	US-09-786-256C-15
16	843.5	21.2	775	4	US-09-786-256C-32
17	764	19.2	812	4	US-09-632-098-4
18	764	19.2	812	4	US-10-177-308-4
19	764	19.2	849	4	US-09-548-797B-6
20	752.5	18.9	606	4	US-09-460-295B-12
21	739	18.5	832	4	US-09-634-252A-4
22	723	18.1	613	4	US-09-026-001A-10
23	721.5	18.1	696	3	US-09-351-414-2
24	721	18.1	814	3	US-09-813-819-4
25	721	18.1	814	3	US-09-920-048-4
26	720	18.1	621	4	US-09-026-001A-18
27	720	18.1	855	3	US-09-813-819-2

28	720	18.1	855	3	US-09-920-048-2	Sequence 2, Appli
29	718	18.0	787	4	US-09-548-797B-5	Sequence 5, Appli
30	717	18.0	621	4	US-09-026-001A-6	Sequence 6, Appli
31	714.5	17.9	571	4	US-09-460-295B-13	Sequence 13, Appli
32	709.5	17.8	616	4	US-09-608-790-1	Sequence 1, Appli
33	701	17.6	802	4	US-09-632-098-2	Sequence 2, Appli
34	701	17.6	802	4	US-10-177-308-2	Sequence 2, Appli
35	700	17.6	769	1	US-08-243-542-4	Sequence 4, Appli
36	700	17.6	769	1	US-08-477-407-4	Sequence 4, Appli
37	700	17.6	769	1	US-08-484-355-4	Sequence 4, Appli
38	698	17.5	670	1	US-08-243-542-3	Sequence 3, Appli
39	698	17.5	670	1	US-08-477-407-3	Sequence 3, Appli
40	698	17.5	670	1	US-08-484-355-3	Sequence 3, Appli
41	698	17.5	867	4	US-09-634-252A-6	Sequence 6, Appli
42	692	17.4	751	2	US-08-836-443-3	Sequence 3, Appli
43	679.5	17.1	592	4	US-09-026-001A-14	Sequence 14, Appli
44	667.5	16.8	746	4	US-09-548-797B-4	Sequence 4, Appli
45	665.5	16.7	540	4	US-09-786-256C-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-08-765-243-8  
; Sequence 8, Application US/08765243  
; Patent No. 5935578  
; GENERAL INFORMATION:  
; APPLICANT: ALVES, KENNETH  
; APPLICANT: GUPTA, SUNIL K.  
; APPLICANT: HOLLIIS, GREGORY F.  
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MARY A. APPOLLINA  
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE  
; CITY: RAHWAY  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/765,243  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: APPOLLINA, MARY A  
; REGISTRATION NUMBER: 34,087  
; REFERENCE/DOCKET NUMBER: 19244Y  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908)594-3462  
; TELEFAX: (908)594-4720  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 734 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-765-243-8

Query Match	100.0%;	Score	3984;	DB	2;	Length	734;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	734;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MWVFLSLGLGRLMDSNFDSLPVQITVPKIRSIKEGIESQASYKIVIEGPKPTVNLN	60				
Db	1	MWVFLSLGLGRLMDSNFDSLPVQITVPKIRSIKEGIESQASYKIVIEGPKPTVNLN	60				
Qy	61	QKNFLPHNFRVYSYSGTGMKPLDQDFNQFNCHYQGYIEGYPKSVVMVSTCTGLRGLQFE	120				

```

Db      61 QKNFLPHNFRVYSYSGTGIMKPLDQFQNFCHYQYIEGYPKSVVMVSTCTGLRGVLQFE 120
QY      121 NVSYGIEPLESSVGFVHHYVQVHKKADVSLYNEKDIESRDLSPKQLQSAEPQODFAKYIE 180
Db      121 NVSYGIEPLESSVGFVHHYVQVHKKADVSLYNEKDIESRDLSPKQLQSAEPQODFAKYIE 180
QY      181 MHVIVEKQLYNHMGSDTTVAQKVFOLIGLTNAIFVSNITIISSLELWIDENKIATTG 240
Db      181 MHVIVEKQLYNHMGSDTTVAQKVFOLIGLTNAIFVSNITIISSLELWIDENKIATTG 240
QY      241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMKCDANYAGGVVHLHPT 300
Db      241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMKCDANYAGGVVHLHPT 300
QY      301 ISLESIAVILAQLLSLNGIYDDINKQCGAGVCIIMNPEAIHFSGVKIFSNCSFEDFAH 360
Db      301 ISLESIAVILAQLLSLNGIYDDINKQCGAGVCIIMNPEAIHFSGVKIFSNCSFEDFAH 360
QY      361 FISKQKSQCLHNPRLDPPFKQAVCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRF 420
Db      361 FISKQKSQCLHNPRLDPPFKQAVCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRF 420
QY      421 KAGSNCAEGPCCECLFMSKERMCRPSFECDLPEYCNCGSSASCENHYYVQTGHPCGLNQ 480
Db      421 KAGSNCAEGPCCECLFMSKERMCRPSFECDLPEYCNCGSSASCENHYYVQTGHPCGLNQ 480
QY      481 WICIDGVCMSGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540
Db      481 WICIDGVCMSGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540
QY      541 CGKLI CKYVGFLLQIPRATIIYANISGHLCTAIEFASDHADSKMKIKDGTSCGSKNVC 600
Db      541 CGKLI CKYVGFLLQIPRATIIYANISGHLCTAIEFASDHADSKMKIKDGTSCGSKNVC 600
QY      601 RNQRCVSSSYLYGDCITTDKNDRGVGNKKHCHCSASYLPDPCSVQSDLWPGGSDSGNF 660
Db      601 RNQRCVSSSYLYGDCITTDKNDRGVGNKKHCHCSASYLPDPCSVQSDLWPGGSDSGNF 660
QY      661 PPVAIPARLPERRYIENIYHSPKMRWPFLLFIPFFIIFCVLIATIMVKNVFORKKWRTEDY 720
Db      661 PPVAIPARLPERRYIENIYHSPKMRWPFLLFIPFFIIFCVLIATIMVKNVFORKKWRTEDY 720
QY      721 SSDEQPESESEPKG 734
Db      721 SSDEQPESESEPKG 734

```

## RESULT 2

```

PCT-US95-07295-8
; Sequence 8, Application PC/TUS9507295
; GENERAL INFORMATION:
; APPLICANT: ALVES, KENNETH
; APPLICANT: GUPTA, SUNIL K.
; APPLICANT: HOLLIS, GREGORY F.
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MARY A. APPOLLINA
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAYWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07295
; FILING DATE:
; CLASSIFICATION:

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: APPOLLINA, MARY A
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19244Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 734 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-07295-8

```

```

Query Match      100.0%; Score 3984; DB 5; Length 734;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MWVLFLLSGLGGRMDSNFDSPVQITVPEKIRSIIEGIESQASVKIVIEGKPYTVNLM 60
Db      1 MWVLFLLSGLGGRMDSNFDSPVQITVPEKIRSIIEGIESQASVKIVIEGKPYTVNLM 60
QY      61 QKNFLPHNFRVYSYSGTGIMKPLDQFQNFCHYQYIEGYPKSVVMVSTCTGLRGVLQFE 120
Db      61 QKNFLPHNFRVYSYSGTGIMKPLDQFQNFCHYQYIEGYPKSVVMVSTCTGLRGVLQFE 120
QY      121 NVSYGIEPLESSVGFVHHYVQVHKKADVSLYNEKDIESRDLSPKQLQSAEPQODFAKYIE 180
Db      121 NVSYGIEPLESSVGFVHHYVQVHKKADVSLYNEKDIESRDLSPKQLQSAEPQODFAKYIE 180
QY      181 MHVIVEKQLYNHMGSDTTVAQKVFOLIGLTNAIFVSNITIISSLELWIDENKIATTG 240
Db      181 MHVIVEKQLYNHMGSDTTVAQKVFOLIGLTNAIFVSNITIISSLELWIDENKIATTG 240
QY      241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMKCDANYAGGVVHLHPT 300
Db      241 EANELLHTFLRWKTSYLVLRPHDVAFLVYREKSNYVGATFQGMKCDANYAGGVVHLHPT 300
QY      301 ISLESIAVILAQLLSLNGIYDDINKQCGAGVCIIMNPEAIHFSGVKIFSNCSFEDFAH 360
Db      301 ISLESIAVILAQLLSLNGIYDDINKQCGAGVCIIMNPEAIHFSGVKIFSNCSFEDFAH 360
QY      361 FISKQKSQCLHNPRLDPPFKQAVCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRF 420
Db      361 FISKQKSQCLHNPRLDPPFKQAVCGNAKLEAGEBCDCGTEQDCALIGETCCDIATCRF 420
QY      421 KAGSNCAEGPCCECLFMSKERMCRPSFECDLPEYCNCGSSASCENHYYVQTGHPCGLNQ 480
Db      421 KAGSNCAEGPCCECLFMSKERMCRPSFECDLPEYCNCGSSASCENHYYVQTGHPCGLNQ 480
QY      481 WICIDGVCMSGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540
Db      481 WICIDGVCMSGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQ 540
QY      541 CGKLI CKYVGFLLQIPRATIIYANISGHLCTAIEFASDHADSKMKIKDGTSCGSKNVC 600
Db      541 CGKLI CKYVGFLLQIPRATIIYANISGHLCTAIEFASDHADSKMKIKDGTSCGSKNVC 600
QY      601 RNQRCVSSSYLYGDCITTDKNDRGVGNKKHCHCSASYLPDPCSVQSDLWPGGSDSGNF 660
Db      601 RNQRCVSSSYLYGDCITTDKNDRGVGNKKHCHCSASYLPDPCSVQSDLWPGGSDSGNF 660
QY      661 PPVAIPARLPERRYIENIYHSPKMRWPFLLFIPFFIIFCVLIATIMVKNVFORKKWRTEDY 720
Db      661 PPVAIPARLPERRYIENIYHSPKMRWPFLLFIPFFIIFCVLIATIMVKNVFORKKWRTEDY 720
QY      721 SSDEQPESESEPKG 734
Db      721 SSDEQPESESEPKG 734

```

RESULT 3

```
US-08-264-101-2
; Sequence 2, Application US/08264101
; Patent No. 5693496
; GENERAL INFORMATION:
; APPLICANT: ALVES, KENNETH
; APPLICANT: GUPTA, SUNIL K.
; APPLICANT: HOLLIS, GREGORY F.
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MARY A. APPOLLINA
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,101
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: APPOLLINA, MARY A.
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-264-101-2

Query Match      89.1%; Score 3551; DB 1; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.3e-284;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      85 QDFQNFCHYQGYEGYKPSVVMYSTCTGLRGVLQFENVSYGIEPLSSVGFHVIYQVKH 144
DB      2 QDFQNFCHYQGYEGYKPSVVMYSTCTGLRGVLQFENVSYGIEPLSSVGFHVIYQVKH 61
QY      145 KKADVSLYNEKDIESRDLSPKLSAEPQDFAKYIEMHVIKQLYNHMGSDTTVVAQKV 204
DB      62 KKADVSLYNEKDIESRDLSPKLSAEPQDFAKYIEMHVIKQLYNHMGSDTTVVAQKV 121
QY      205 FQILGLTNALFVSNITILSSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 264
DB      122 FQILGLTNALFVSNITILSSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 181
QY      265 AFLVVRKSNYYGATFGQKWCNANVAGGVVLTPTLSLESLAVILAQLLSLNGITYDD 324
DB      182 AFLVVRKSNYYGATFGQKWCNANVAGGVVLTPTLSLESLAVILAQLLSLNGITYDD 241
QY      325 INKQCQSGAVCINPBAIHFSGVKIFSNCSFEDFAHFIKQSKQCLHNQRLDPFFKQQA 384
DB      242 INKQCQSGAVCINPBAIHFSGVKIFSNCSFEDFAHFIKQSKQCLHNQRLDPFFKQQA 301
QY      385 VCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGNCNCAEGPCCNCLFMSKERM 444
DB      302 VCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGNCNCAEGPCCNCLFMSKERM 361
QY      445 RPSFECDLPEYCNCGSSACPNENHYVTGHPCCGLNOMICDGVCMGDKQCTDTFGKEVE 504
DB      362 RPSFECDLPEYCNCGSSACPNENHYVTGHPCCGLNOMICDGVCMGDKQCTDTFGKEVE 421

US-08-765-243-2
; Sequence 2, Application US/08765243
; Patent No. 5935578
; GENERAL INFORMATION:
; APPLICANT: ALVES, KENNETH
; APPLICANT: GUPTA, SUNIL K.
; APPLICANT: HOLLIS, GREGORY F.
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MARY A. APPOLLINA
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,243
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: APPOLLINA, MARY A.
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19244Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-765-243-2

Query Match      89.1%; Score 3551; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.3e-284;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      85 QDFQNFCHYQGYEGYKPSVVMYSTCTGLRGVLQFENVSYGIEPLSSVGFHVIYQVKH 144
DB      2 QDFQNFCHYQGYEGYKPSVVMYSTCTGLRGVLQFENVSYGIEPLSSVGFHVIYQVKH 61
QY      145 KKADVSLYNEKDIESRDLSPKLSAEPQDFAKYIEMHVIKQLYNHMGSDTTVVAQKV 204
DB      62 KKADVSLYNEKDIESRDLSPKLSAEPQDFAKYIEMHVIKQLYNHMGSDTTVVAQKV 121
QY      205 FQILGLTNALFVSNITILSSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 264
```

Db 122 FQILGLTNAIFVSNITILSSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 181  
Qy 265 AFLVYREKSNVYGATFGKMCNDANYAGVVLHPTTISLESILAVILAQLLSLMSGITYDD 324  
Db 182 AFLVYREKSNVYGATFGKMCNDANYAGVVLHPTTISLESILAVILAQLLSLMSGITYDD 241  
Qy 325 INKQCSCGAVCIMNPEAIHFSGVKIFSNCSFEDFAHFIKSKQSCCLHNQRLDPFFKQQA 384  
Db 242 INKQCSCGAVCIMNPEAIHFSGVKIFSNCSFEDFAHFIKSKQSCCLHNQRLDPFFKQQA 301  
Qy 385 VCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERM 444  
Db 302 VCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERM 361  
Qy 445 RPSFECDLPEYCNCGSSASCENHYVQTGHPCGLNQWICIDGVCMGSDGKQCTDTFGKEVE 504  
Db 362 RPSFECDLPEYCNCGSSASCENHYVQTGHPCGLNQWICIDGVCMGSDGKQCTDTFGKEVE 421  
Qy 505 FGPSECVSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLI CKYVGFLLQIPRATIYA 564  
Db 422 FGPSECVSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLI CKYVGFLLQIPRATIYA 481  
Qy 565 NISGHLCIAVEFASDHADSOQWIKDGTSCGSKNVCNRQCVSSSYLGIDCTTDKCDNRG 624  
Db 482 NISGHLCIAVEFASDHADSOQWIKDGTSCGSKNVCNRQCVSSSYLGIDCTTDKCDNRG 541  
Qy 625 VCNKXKHCHCSASYLPDSCVQSDLWPGGSDSGNFPFPAIPARLPERRRYENIYHSPKM 684  
Db 542 VCNKXKHCHCSASYLPDSCVQSDLWPGGSDSGNFPFPAIPARLPERRRYENIYHSPKM 601  
Qy 685 RWPFFLFIPIFPCVLIAIMVKNFQKWKRTEDYSSDEQPESEPKG 734  
Db 602 RWPFFLFIPIFPCVLIAIMVKNFQKWKRTEDYSSDEQPESEPKG 651

## RESULT 5

PCT-US95-07295-2  
; Sequence 2, Application PC/TUS9507295  
; GENERAL INFORMATION:  
; APPLICANT: ALVES, KENNETH  
; APPLICANT: GUPTA, SUNIL K.  
; APPLICANT: HOLLISS, GREGORY F.  
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MARY A. APOLLINA  
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE  
; CITY: RAHWAY  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/07295  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: APOLLINA, MARY A  
; REGISTRATION NUMBER: 34,087  
; REFERENCE/DOCKET NUMBER: 19244Y  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908)594-3462  
; TELEFAX: (908)594-4720  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 651 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
PCT-US95-07295-2

Query Match 89.1%; Score 3551; DB 5; Length 651;  
Best Local Similarity 100.0%; Pred. No. 4.3e-284; Indels 0; Gaps 0;  
Matches 650; Conservative 0; Mismatches 0;

Qy 85 QDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVIYQVGH 144  
Db 2 QDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFHEVIYQVGH 61  
Qy 145 KKADVSLYNEKDIIESRDLSPKLSABPQDFAKYIEMHVIIVEKOLYNHMGSDTTVAQKV 204  
Db 62 KKADVSLYNEKDIIESRDLSPKLSABPQDFAKYIEMHVIIVEKOLYNHMGSDTTVAQKV 121  
Qy 205 FQILGLTNAIFVSNITILSSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 264  
Db 122 FQILGLTNAIFVSNITILSSLELWIDENKIATTEANELLHTFLRWKTSYLVLRPHDV 181  
Qy 265 AFLVYREKSNVYGATFGKMCNDANYAGVVLHPTTISLESILAVILAQLLSLMSGITYDD 324  
Db 182 AFLVYREKSNVYGATFGKMCNDANYAGVVLHPTTISLESILAVILAQLLSLMSGITYDD 241  
Qy 325 INKQCSCGAVCIMNPEAIHFSGVKIFSNCSFEDFAHFIKSKQSCCLHNQRLDPFFKQQA 384  
Db 242 INKQCSCGAVCIMNPEAIHFSGVKIFSNCSFEDFAHFIKSKQSCCLHNQRLDPFFKQQA 301  
Qy 385 VCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERM 444  
Db 302 VCGNAKLEAGEECDCGTEQDCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERM 361  
Qy 445 RPSFECDLPEYCNCGSSASCENHYVQTGHPCGLNQWICIDGVCMGSDGKQCTDTFGKEVE 504  
Db 362 RPSFECDLPEYCNCGSSASCENHYVQTGHPCGLNQWICIDGVCMGSDGKQCTDTFGKEVE 421  
Qy 505 FGPSECVSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLI CKYVGFLLQIPRATIYA 564  
Db 422 FGPSECVSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLI CKYVGFLLQIPRATIYA 481  
Qy 565 NISGHLCIAVEFASDHADSOQWIKDGTSCGSKNVCNRQCVSSSYLGIDCTTDKCDNRG 624  
Db 482 NISGHLCIAVEFASDHADSOQWIKDGTSCGSKNVCNRQCVSSSYLGIDCTTDKCDNRG 541  
Qy 625 VCNKXKHCHCSASYLPDSCVQSDLWPGGSDSGNFPFPAIPARLPERRRYENIYHSPKM 684  
Db 542 VCNKXKHCHCSASYLPDSCVQSDLWPGGSDSGNFPFPAIPARLPERRRYENIYHSPKM 601  
Qy 685 RWPFFLFIPIFPCVLIAIMVKNFQKWKRTEDYSSDEQPESEPKG 734  
Db 602 RWPFFLFIPIFPCVLIAIMVKNFQKWKRTEDYSSDEQPESEPKG 651

## RESULT 6

US-08-765-243-6  
; Sequence 6, Application US/08765243  
; Patent No. 5935578  
; GENERAL INFORMATION:  
; APPLICANT: ALVES, KENNETH  
; APPLICANT: GUPTA, SUNIL K.  
; APPLICANT: HOLLISS, GREGORY F.  
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MARY A. APOLLINA  
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE  
; CITY: RAHWAY  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS



```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,243
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: APPOLLINA, MARY A
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19244Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-765-243-6

Query Match      61.2%; Score 2440; DB 2; Length 735;
Best Local Similarity 59.6%; Pred. No. 1.9e-194;
Matches 439; Conservative 119; Mismatches 167; Indels 12; Gaps 717

Qy      3 VLFLLSG---LGLL---RMDSNFDSLPLVQIITVPEKIRSIIEKEIESQASYKIVIEGPKYT 56
Db      4 ILLLLSELGSELGSSQSQTEGTEKRLHVQVTVPEKIRSVTSNGYETQVTYNLKIEGKTYT 63

Qy     57 VNLQMKFLPHNFRVYSVSGTGIMKPLDDQDFNQFCHYOGYIEGYPKSVVMYSTCTGLRGV 116
Db     64 LDLQMKFLPPNFRVYSVDNAGIWRSLUEKQFNICYFOGYIEGYPNSVMVYSTCTGLRGF 123

Qy    117 LQFENVSYGIEPLESSGFHVIVQVHKHKADVSLYNEKDIIESRDLSPKLOASFPQDDFA 176
Db    124 LQFENVSYGIEPLESSGFHVIVQVPEKGGALLYAEKDIIDRLDSQYKIRSIKQRIVS 183

Qy    177 KYIEMWVIVEKQLYNHNGSDTTVVAQKVFLQIGLTNAIFVFSNITITLSSLELMIDENKI 236
Db    184 HYLEIHTVWEKQMFHEIGADTAIVTQKIFQLIGLANAIFAPFNLTITLSSLEFWMDENKI 243

Qy    237 ATTCEANELLHTLRWKTSLYLVRPHDVALLVYREKSNVYGATFOGKMCADANYAGGVVL 296
Db    244 LTTGDANKLLRFKWKQSLYLVRPHDMAFLLVYRNTTDDYVGATYQGMCKDKNYAGGVAL 303

Qy    297 HPRTISLESVILAQLLSLSMGITYDDINKCQCSGAVCINMNPBAIHFSGVKIFSNCSFE 356
Db    304 HPKAVLTESLAILVLQLLSLSMGLAYDDVNVKCCQGVPCVWNPAPHSSGVRAFSNCSME 363

Qy    357 DPAHFISKQSLCHLNQRPRLDPFPKQQAQVGNAKLEAGEBCDCGTGTEQDICALIGTCCDIA 416
Db    364 DFSKFIITSQSHCLQNQPTLQPSYK-WAVCCNGEVEDEICDCG-KKGCABMPPPCCNPD 421

Qy    417 TCRPKAGSNCAEGGCCENCLFMSKERWCRPSFECDLPEYCNGSGASCPENHYVQTGHPC 476
Db    422 TCKLSDSGSESSGICCNCKLKRKGEVCRLAQDECDVTEYCNGTSEVC-EDFFVQNGHPC 480

Qy    477 GLNOWICIDGVMSSGDQCTDTFGKEVEFGPSECVSHLNSKTDVSGNCSGSDSGVTOCEA 536
Db    481 DNRKWCINGTCQSGEQCCQDLFGIDAGFGSSECFWELNSKSDISGSGCISAGYKECP 540

Qy    537 DNLQCGKLICKYVGKFLQIPRATIIYANISGHILCIAVEFASDHADSQRMWKIDGTSCGS 596
Db    541 NDRMCGKIICKYQSENILKLRSAIVIANISGHVCSVLEYPQGHNESQKMWVRDGTVCGS 600

Qy    597 NKVCNRQRCVSSYLVGDCITTDKNDRGVCMNKKHCHCSASYLPRDCSVQSDLPFGGSID 656
Db    601 NKVCNQKCVADTFLGVDNLEKCNHHGVCNCKNCHCDPTFLPDPCKRMKDSYFGSID 660

Qy    657 SGNPPPPVAILPARLPERRYENIYHSPMRWPFPLFIPFIIIFCVLIIMVKVNFQRKKWR 716
Db    661 SGN-KERAEP--IPVRPIIASRYKSGRPWPFLLIIPFYVILVLIGMLVKVYSQRMKWR 717

Qy    717 TEDYSDEQPESESEPK 733

```

Db 364 DFKSKFITSOSSHCLQNOPTLQPSYK-MAVCGNGEVEDEICDGG-KKGCACEMPPCCNPD 421  
Qy 417 TCRFKAGSNCAEGPCENCLFMSKERMCRPSFECDLPEYCGSSASCPENHYVOTGHPC 476  
Db 422 TCKLSGSECSGICGNCCKLKRKGEVCRUQADECVDTEYCNCTSEVC-EDFFVQNGHPC 480  
Qy 477 GLNQWICIDGVCMGDKQCTDTDFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEA 536  
Db 481 DNRKWCINGTCGSEGOCCDLFGIDAGFGSSCFWELNSKSDISGSCGISAGYKECPP 540  
Qy 537 DNIQCKGLICKYVGKFLQIPRATIIYANISGHLCTIAVEFASDHADSOQXWIKDGTSCGS 596  
Db 541 NDRMCGKIIKCYQSENILKLSATVIYANISGHVCSLEYPGQHNSQKMWVRDGTVCGS 600  
Qy 597 NKVCRNORCVSSSYLVGDCCTDKNDRGVGNKKGCHCSASYLPDCCSVQSDLWPGSID 656  
Db 601 NKVQCNQKQVADTFGLGYDCNLEKCNHGVGNKKNCHCDPTLYLPDCKRMKDSYPGGSID 660  
Qy 657 SGNFPPVAIPARLPERRYENIYHSPKMRPFFFLFIPFFIIFCVLAIAMVKNVFORKKWR 716  
Db 661 SGN-KERAEP--IPVRPIASRYRSKSPRPWFFFLIIPFVVVILVILGMLVKVYSQRMKWR 717  
Qy 717 TEDYSDEQSESESEK 733  
Db 718 MDDFSSEBEQSESESEK 734

## RESULT 8

US-08-264-101-4  
; Sequence 4, Application US/08264101  
; Patent No. 5693496  
; GENERAL INFORMATION:  
; APPLICANT: ALVES, KENNETH  
; APPLICANT: GUPTA, SUNIL K.  
; APPLICANT: HOLLIS, GREGORY F.  
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: MARY A. APPOLLINA  
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE  
; CITY: RAHWAY  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/264,101  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: APPOLLINA, MARY A.  
; REGISTRATION NUMBER: 34,087  
; REFERENCE/DOCKET NUMBER: 19244  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908)594-3462  
; TELEFAX: (908)594-4720  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 457 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-264-101-4

Query Match 38.0%; Score 1512; DB 1; Length 457;  
Best Local Similarity 57.6%; Pred. No. 3.1e-116;  
Matches 265; Conservative 74; Mismatches 115; Indels 6; Gaps 5;

Qy 274 SNYVGATFGKMCMDANYAGGVVLHPRITISLESIAVLIAQLLSMGITVDDINKKCCSGA 333  
Db 3 SDYVGATYQKCMCDKNYAGGVVALHPKAVTLESIAIILVQLLSMLGLAVDDVNNKCCQGV 62  
Qy 334 VICMPEAIFSGVKIFSNCSFEDFAHFTSKQKSOCLHNPRLDPFFKQOAVCGNAKLEA 393  
Db 63 VCMNPEAPHSGVRAFSNCSMEDFSKFTSOSSHCLQNOPTLQPSYK-MAVCGNGEVEE 121  
Qy 394 GBECDGTQDCCALIGETCCDIATCRFKAGSNCAEGPCENCLFMSKERMCRPSFECDL 453  
Db 122 DEICDGG-KKGCACEMPPCCNPDTCCKLSGSECSGICGNCCKLKRKGEVCRUQADEC 180  
Qy 454 PEYCGSSASCPENHYVOTGHPCGLNQWICIDGVCMGDKQCTDTDFGKEVEFGPSECYSH 513  
Db 181 TEYCNCTSEVC-EDFFVQNGHPCDNRKWCINGTCGSEGOCCDLFGIDAGFGSSECFWE 239  
Qy 514 LNSKTDVSGNCGISDSGYTQCEADNLQCKLICKYVGKFLQIPRATIIYANISGHLCTA 573  
Db 240 LNSKSDISGSCGISAGYKECPPNDRMCKIIICKYQSENILKLSATVIYANISGHVCS 299  
Qy 574 VEFASDHADSOQXWIKDGTSCGSKNVCNRQRCVSSSYLVGDCCTDKNDRGVGNKKGCH 633  
Db 300 LEYPOGHNESQKMWVRDGTVCNKNVCQKQCVADTFGLGYDCNLEKCNHGVGNKKNCH 359  
Qy 634 CSASYLPDCCSVQSDLWPGGSIDSGNFPFVAIPARLPERRYENIYHSPKMRPFFFLFIP 693  
Db 360 CDPTLYLPDCKRMKDSYPGGSIDSGN-KERAEP--IPVRPIASRYRSKSPRPWFFFLIIP 416  
Qy 694 FFIIFCVLAIAMVKNVFORKKWRTEDEYSDDEQSESESEK 733  
Db 417 FVVVILVILGMLVKVYSQRMKWRMDDFSSEBEQSESESEK 456

## RESULT 9

US-08-765-243-4  
; Sequence 4, Application US/08765243  
; Patent No. 5935578  
; GENERAL INFORMATION:  
; APPLICANT: ALVES, KENNETH  
; APPLICANT: GUPTA, SUNIL K.  
; APPLICANT: HOLLIS, GREGORY F.  
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: MARY A. APPOLLINA  
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE  
; CITY: RAHWAY  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/765,243  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: APPOLLINA, MARY A.  
; REGISTRATION NUMBER: 34,087  
; REFERENCE/DOCKET NUMBER: 19244  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908)594-3462  
; TELEFAX: (908)594-4720  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 457 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

```

; MOLECULE TYPE: protein
US-08-765-243-4

Query Match      38.0%; Score 1512; DB 2; Length 457;
Best Local Similarity 57.6%; Pred. No. 3.1e-116;
Matches 265; Conservative 74; Mismatches 115; Indels 6; Gaps 5;

Qy 274 SNTVGATFGKMCNDANYAGGVVLPRTTISLESIAVLQAQLLSLGMITYDDINKCOCGSA 333
Db 3 SDYVGATYGCKMDKNYAGVALHPKAVTLSTAILVQLLSMLGLAYDDVNKCQCQGPV 62

Qy 334 VCIWNPFAHFSKVIFSNCSPEDEAFHFISKOKSOCLHNOPRLDPFFKQOAVCGNAKLEA 393
Db 63 VCWNPEAPHSSGVRAPFNSCMEDFSKFITSOSSHLQNQPTLQPSYK--MAVCGNGVEE 121

Qy 394 GEBCDCGTETQDCCALIGETCCDIATCFRAGSNCABGPCCECNLFMSKERMCRPSPPEECDL 453
Db 122 DEICDCG-KKGCAEMPPPCNPDTCKLSDGSECSGGICCNCKLRKEVCVELAQDECVDV 180

Qy 454 PEYCNGSSASCPENHVQTGHPCGLNQMTCIDGVCMGSKQCTDTFGKEVFEPGFPSECYSH 513
Db 181 TEYCNGTSEVC-BDFVQNHGPHCDNRKWICINGTCOSGEQQDLFLGDAGFSSECFWE 239

Qy 514 LNSKTDSVSGNGISDSGYTQCEADNLQCKLKICKYVKFKLLQIPRATIYANISHGLCIA 573
Db 240 LNSKSDISGSCGISAGGYKECPNDRMCKGIICKYQSENILKLRSATVIYANISHGVCS 299

Qy 574 VEFASDHADSQKMWKIDGTSCGSNKVCRNQRVSSSYLYGDCYTTCDCNDRGVGNKKHCH 633
Db 300 LEYPQHNSQKMWVRDGTVCASNKVCQKQCVATFLGYDCNLEKHNGHGVGNKKHCH 359

Qy 634 CSASYLPDPCSVDLWPGSIDSGNFPPVAIPARLPERRYIENTIYHSKPMRWPFPLPIP 693
Db 360 CDPTYLPDCKRMKDSYPGSIDSGN-KERAEP--IPVRPIASRYRSKSPRPWPFLIIP 416

Qy 694 FFIIFCVLIAMVKNFORKKWRTEYSSDBQESESSEPK 733
Db 417 FYVIVLVLMGLVKVYSQRMKWMDDFSSEEQFESESESK 456

RESULT 10
PCT-US95-07295-4
; Sequence 4, Application PC/TUS9507295
; GENERAL INFORMATION:
; APPLICANT: ALVES, KENNETH
; APPLICANT: GUPTA, SUNIL K.
; APPLICANT: HOLLIS, GREGORY F.
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MARY A. APPOLLINA
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07295
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: APPOLLINA, MARY A
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19244Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; INFORMATION FOR SEQ ID NO: 4:

```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..16
US-10-000-489-70

Query Match      37.8%; Score 1506.5; DB 4; Length 787;
Best Local Similarity 40.7%; Pred. No. 1.9e-115;
Matches 307; Conservative 135; Mismatches 261; Indels 51; Gaps 18;

QY 1 MWVLFLLSGGLRMD-SNFDLSLPVQITVPEKIRSIIEKIEGIE---SQASQKIVIEGPKYT 56
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 LWL--LLAGCGLLASPGFNSLLQIVPEKIQTNTDSSEIEYEQISYIIPIDEKLYT 61
QY 57 VNLQKQFLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQGYIEGPKSVVMVSTCTGLRGV 116
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 VHLKQRYFLDNFWIYLYN-QGSNWTYSDDIQCYIQGNIEEYPDSNMTLSICSLRGVI 120
QY 117 LQFENVSYGIEPLESSVGFPHVIYQVHKKADVSLYNEKDIESR---DLSFKLQSAEPQ 173
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 LQFENVSYGIEPLESAVEFQVHLKLVNDNDIAIFDRSLKEQPMDDNIFISEKSEPAV 180
QY 174 D--FAKIEHVIIVEKOLYNHMGSDTTVAQKVPQLIGLNTAIFVSNFIILISLELWI 231
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 PDLFPLYLEHVIYVDKTLTYDYGSDSDMIVTNKVIIEIVGLANSMTQPKVTIVLSLELWS 240
QY 232 DENKIATGBANELLHTFLRWKTSYVLRPHDVAFLVYREKSNYVGATQGMKCDANYA 291
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 DENKISTVGADELLQKFLWKQSYLNLPHDIAIYLYMDYPRYLGAVPFGTWCITRYS 300
QY 292 GGVLHPRTISLESIAVILAQLLSLGMITYDDINKCQCGSAGVCMNPEAHIFSGVKIFS 351
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 AGVALYKPEITLFAFVIVTQMLSLGISYDDPKQCQSESTCINNPEVQSNVGVKTF 360
QY 352 NCSFEDPAHISKQSOCLHNQPLDPFFKQQAQVCGNAKLEAGECDCEGTEQDCALIGET 411
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361 SCSLRSQNFISNVGVKCLQNKPMQK-KSPKPYCGNGRLGNEICDCGTEAQCG--PAS 417
QY 412 CCDIATCRKAGSNCAEGPCENCLFMSKERMCRP-SFEEDCLPEYCNAGSSASCPENHYV 470
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
418 CDFRITCVLDGAKYKGLCKDCQIILQSGVECPKAPHAPECDIAENCGSSPCEGPDITL 477
QY 471 QTGHPGCLNQMICIDGVCMSGDQKQCTDTFGKEVEFGPSEYSHLNSKTDVSGMCGIS-DS 529
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
478 INGLSCKNNKFIYDGDCHLDARCESVFGKGRNAPFACVEEIQSQSDRFGNCGDRNN 537
QY 530 GYTQCEADNLCQGLIKYVCKFLQIIPRATIIYANISGHLCIAVEPASDHADSQKWKIK 589
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
538 KYVFCGRNRLICGLRVCTYPRKPFHQENGVDVIAFVRDVSVCITVDYKLPRTVPDPLAVK 597
QY 590 DGTSCGSKNVCNRCNORCVSSSYL---GYDCTDKNDRGVNCKKHCHCSASYLPDQCSVQ 646
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
598 NGSQCDIGRCVNVRECVESRIIKAHVC-SQQCSGHGVCDNRNKHCHCSGYKPPNCQIR 656
QY 647 SLDWPGGSI---DSGNFPFVPAIPARLPERRYIENIYHSPMRW--PFFIIPFPIIFCV 700
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
657 S---KGFSIPPEDMGSI-----MERASGKTENTWLLGLFIALPILIV---696
QY 701 LIAIMVKNVORKWRTEDYSSDRQPESEPKG 734
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
697 -TTAIVLARKQLKNWFAKE---EEFSPSEKSEG 726

RESULT 12
US-10-140-002-116
; Sequence 116, Application US/10140002
; Patent No. 6725730
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
```

QY 650 WPGSIDSNGNPPVAIPARLPERRIENIYHSEKMRW---PFLFPIPFPIFCVLIAMV 706  
Db 636 -PGSIDDGNGF-----QKSGDFYTEKGYNTHWNWFIILSCIFLFPFVFTVI---F 684  
QY 707 KVNFORKWRTE--DYSSDQSESESPKG 734  
Db 685 KRNEISKCRNENAYNRNNSVSESDVG 714

## RESULT 13

US-10-140-002-204  
; Sequence 204, Application US/10140002  
; Patent No. 6725730  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C59  
; CURRENT APPLICATION NUMBER: US/10/140.002  
; CURRENT FILING DATE: 2002-05-06  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 204  
; LENGTH: 790  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-140-002-204

Query Match 23.8%; Score 948.5; DB 4; Length 790;  
Best Local Similarity 32.0%; Pred. No. 1.8e-69;  
Matches 254; Conservative 124; Mismatches 292; Indels 123; Gaps 30;

QY 6 LLSGLGG---LRMDSNFDLPVQITVPEKI--RSIIKEGIESQASYKIVIEGKPYTVNLM 60  
Db 22 LLKSLGEDVIFHPGEFDS--YEVTIPEKLSFRGEV-QGVVSPVSYLLQLKGGKHVLHW 78  
QY 61 QKN-FLPHNFRVYSYSGTGIMKPLDQDFQNFCHYQVIEGPKSVVMVSTCT-GLRVQL 118  
Db 79 PKRLPLRLRVSFTEHGELEDDHPYIPKDCNMGVSKESLDSKATISTCMGRLRGVFN 138  
QY 119 FENVSYGIEPLESSVGFEHVIYQVKKKADVSLYNEKDIERSDLSFKLQSAEPQDF--- 175  
Db 139 IDAKHQVIEPLKASPSFEHVVYLLKKEQFGNQVCGLSDDSEIEMQAPYENKARLURDPGS 198  
QY 176 ---AKYIEMHIVIEKQLYNHMGSDTTVVAQKVFQILGTLNATFVSNITILLSSLELWID 232  
Db 199 YKHPKYLELILLFDQSYRFPVNNLSQVTHDAILLTGIMDTYFQDVERMRTHLKALEWTD 258  
QY 233 ENKIAT--TGANELLHTFLRWKTSYLVLR-PhdVAFLLVYREKSNVYVATFGQKMDANY 290  
Db 259 FNKIRVGYPELAEVLGRFVIYKKSVNLNARLSSDWAHLQKRYNDALAWSF-GKVSLEY 317  
QY 291 AGGV-----VLHPTISLESVALAQLLSLMSGTYDDINKCQCSGAV-CIMNPEAI 342  
Db 318 AGSVSTLLDNTILAPATWS-----AHELGHAVGMSHDE-QYCQCRGLNLCIMG---- 364  
QY 343 HFSGVKIFSCSPEDFAHFTISKQSKQLCHNQPLRDPFFKQAVCGNKALEAGEBCDCGTE 402

Db 365 --SGRTGFSNCSYISFFKHIS-SGATCLNNIPGLGYVLKR---CGNKIVEDNEEDCCGST 418  
QY 403 QDCALLGETCCDIATCRFKAGSNCAEGPCCECNCLFMSKERMCMRSPSPBECOLPEYCNSSA 462  
Db 419 EEQ--KORCCQ-SNCKLQFGANCSIGLCCDCHFRPSGYVCRQEGNECDLAECYDGNSS 475  
QY 463 SCPENHYVQTHGPCGLNQWICIDGVCMSGDKQCTDTFTGKEVFEFSPSCYSHLNSKTDVSG 522  
Db 476 SCPNDVYQDGTCKY-EGRCFRKCRSRMYQCQSIQFPDAMEAPSECYDAVNLIQDQFG 534  
QY 523 NCGISD-SGYTQCEADNLQCKLICKYVGFELQIIPRATIIYANISHLCAIEFADHA 581  
Db 535 NCEITGIRNFKCESANSICGRLOQINV-ETIPDLPEHTTI---ISFHL-----QA 581  
QY 582 DSQKMW-----IKDGTSCGNKVCNRQCRVSSSYLVGYDCTTDCNKDR 623  
Db 582 ENLMCWGTGYHLSMKPMGIPDLGMINDTSCGSGRVCFKNCVNSVLPQDCLPEKCNTR 641  
QY 624 GVCNNKXKHCASAYLPDCSVQSDLWPGSIDSGNFPFVAIPARLPERRIENIYHSPK 683  
Db 642 GVCNNRKNCHCMYGWAPPFC---EEVGYGGSIDSG--PPGLLRGAIPSSIWVSI----- 691  
QY 684 MRWPFPLPIPFPIFCVLIAMV-----KVNFORKKWRT-EDY 720  
Db 692 -----IMFRLILLISVYVFFRQVIGNHLKPKQEKMLPSKAKTEQESKTKTVQEE 743  
QY 721 SSDEQSESESPK 733  
Db 744 SKYTKQEESEAK 756

## RESULT 14

US-09-617-145-2  
; Sequence 2, Application US/09617145  
; Patent No. 6485956  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: SVPH1-8 DNA and Polypeptides  
; FILE REFERENCE: 03260.0050-00304  
; CURRENT APPLICATION NUMBER: US/09/617,145  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/071,505  
; PRIOR FILING DATE: 1998-01-14  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 722  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-617-145-2

Query Match 23.5%; Score 937; DB 4; Length 722;  
Best Local Similarity 31.0%; Pred. No. 1.4e-68;  
Matches 219; Conservative 133; Mismatches 284; Indels 70; Gaps 24;

QY 1 MWVLLSLGGLGGMRMDSNFDLPVQITVPEKIRSIKEGIESQA---SYKIVIEGKPYT 56  
Db 17 LNLGVFLSISGYCQAGPSQHTSPFVVIPUK---VLSRGRSAKAPGWLSTLRFGGQKHV 73  
QY 57 VNLQMKNFL-PhnFRVYVSYSGTGMKPLDQDF-QNFCHYQVIEGPKSVVMVSTC-TGL 113  
Db 74 VHMVRVKLLVSRHLPVFTYTDERALLE-DQLFIPDDCYHYGVGAPESLVVPSACFGGF 132  
QY 114 RGVLFQENVSYGIEPLESSVGFEHVIYQVKKKADVSL-----YNEKDIERSDLSF---KL 166  
Db 133 RGVLKISGLTYEISPIRHSATFELHVVYKNSNETQFPAMRCGLTEKEVARQQLFEBAEN 192  
QY 167 QSAEPQ-----DFAKYIEMHIVIEKQLYNHMGSDTTVVAQKVFQILGTLNATFVSNIT 221  
Db 193 SALEPESAGDWTWTHAFLELVVNVNHDFFIYSQNSISKVQEDVFLVNVNIQQQLGTY 252  
QY 222 ILLSLELWIDENKIATTTGANELLHTFLRWKTSYLVLRPHDVAFLLVYREKSNVYVATG 281

```

Db 253 IILIGBIWQGNVFPMT-SIEQVINDFSQWKQISLSQLQDHAHMFKNLSLSILGLAY 311
QY 282 QGKMCDA-----NYAGGVVHLPTISLESIAVLAQLLSLMSGITDYDINKQCGSGAV 334
Db 312 VAGICRPPICGVDNFOGD-----TWSL--FANTVAHELCHTLGMOHDE-EFCFCGERG 362
QY 335 CIMNPEAHFSGVKIFGSCFEDFAHFISSKQSOCLNQPRLDPPFFQQAQVCGNAKLEAG 394
Db 363 CIMNTPRV--PAEKFTNCVADPMKTTNLQGS-CLNPPRLGIFMLKR-CGNGVVERE 417
QY 395 BECDGTEQCALIGETCCDIATCRFAGSNCARGPCENCLFMSKERMCRPSFEBCDLP 454
Db 418 EQCDGSGVOQCE--QDACC-LLNCTLRPGAACAFGLCKCKCKFMPSELCRQEVNECDLP 474
QY 455 EYCNNGSASCPENHYVOTGHPGCLNQWICIDGVCMGDKOCTDFGKEVEFGFSECYSHL 514
Db 475 EWNGTSHQCPEDRYVQDGIPCSQAY-CYQKRCNNHDDHCRREIFGDKASQNCYKEI 533
QY 515 NSKTDVSGNGCISDSGYTOCEADNLOCGKLYCKYVGRFLQIPRATIIYANISGHLCAV 574
Db 534 NSQNRFGHCGINGTTLYLKCHISDFVCGRVOCENVRDIPLLQDHFTHLQTHINGVTCWI 593
QY 575 EF-----ASHADSQKWIIDGTCGSKNKKVCRNORCVSSSYLVGDCTTDKNDRGVGNK 629
Db 594 DYHLRMNISDGE-----VKDGTVCGPCKICHHKCVSLSVLSHVCLPETCNMKGICNNK 648
QY 630 KHCHCSASYLPDPCSVQSDLWPGGSDSGN-----FPPVAIPA 667
Db 649 HHCHCGXWSPYQCQHRG---YGSIDSGPASAKRGVFLPLIVIPS 691

RESULT 15
US-09-786-256C-15
; Sequence 15, Application US/09786256C
; Patent No. 6680189
; GENERAL INFORMATION:
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: HIKICHI, Yuichi
; APPLICANT: NISHIMURA, Atsushi
; TITLE OF INVENTION: No. 6680189el Protein and DNA Thereof
; FILE REFERENCE: 2544 USOP
; CURRENT APPLICATION NUMBER: US/09/786,256C
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: PCT/JP99/04766
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: JP 10-250115
; PRIOR FILING DATE: 1998-09-03
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 15
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(775)
; OTHER INFORMATION: An isolated ADAM family protein
US-09-786-256C-15

```

```

Query Match 21.2%; Score 843.5; DB 4; Length 775;
Best Local Similarity 28.9%; Pred. No. 8.1e-61;
Matches 210; Conservative 121; Mismatches 300; Indels 95; Gaps 23;

QY 29 PEKIRSIIEKESQASQYKIVIEGKPYTVNLQ-KNPLPHNFRVYSYSGTGMKPLDQDF 87
Db 47 PEQ-----QEQFETELKYKMTINGKIAVLKKNKLLAPGYTETYNSTGKEITTSPOI 101
QY 88 QNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFENVSYGIEPLE--SSVGFEHVI--YQVK 143
Db 102 MDDCYQGHILNEKVSQASISTCGLRGYSQDQRYFIEPLSPIHRDQGEHALFKYNPD 161
QY 144 HKKADVSLYNEKQIESRDLs-----FKLQSAEPQODFAKYIEMHVIVEKQLYNHM 193

```

```

Db 162 EKNYDSTCGMDGVLWADLQONIALPATKLVKLKDKRK-VQEHEKYIEYLVLDNGEPKRY 220
QY 194 GSOTTVVAQKVFOLIGLTNAIFVSFNITITLSLELWIDENKIATTGEANELLHTFLRWK 253
Db 221 NENQDEIRKRVFEMANVVMYKLLTHVALVGMETWTDKDKIKITPNASFTLENFSKWR 280
QY 254 TSYLVLRP-HDVAFLVLYRE-KSNYYGATFQGMKCDANYAGGVVLPRTISLESIAVLA 311
Db 281 GSVLSRRKGRHDAQLTATATELAGTTVGLAFMTMC-SPYSVGVV-QDHSNLLRLVAGTMA 338
QY 312 QLLSLSMGITYDDINKQCGSAGVCIWNPRAIHFSGVKIFSNCSFEDFAHFISSKQSOCLH 371
Db 339 HEMGHNFMPHDDYS-CKPSTTICVMD-KALSFYIPTDFSSCSRLSYDKFEFDKLSNCLF 396
QY 372 NQRLDPFFQQAQVCGNAKLEAGEECDCGTEQCALIGETCCDIATCRFAGSNCARGPC 431
Db 397 NAP-LPTDIIISTICGNQLVEMGEDCCTGSECTWI---CCDAKTKIKATFQCALGEC 452
QY 432 CENCLFMSKERMCRPSFEBCDLPPEYCNNGSASCPENHYVOTGHPGCLNQWICIDGVCMG 491
Db 453 CEKQCFKAGMVCRAKDECDLPEMCMGKSGNCPDORFQVNGFPFCHHGKGLMGTCPIL 512
QY 492 DKQCTDTFGKEVEFGFSECYSHLSKTDVSGNGCISDSGYTOCEADNLOCGKLYCK 547
Db 513 QEQCTELWGPTEVADKSCYNR-NEGGSYGYCRRVDDTLIPCKANDTMCGLKLPFCOGSD 571
QY 548 ---YVGKFLQIPRATIIYANISGHLCAVEFASDHADSQKWIIDGTCGSKNKKVCRNOR 604
Db 572 NLPWKGRIVTFLTKT-----FPEDTSQEIGMVGANGTKCDNKKVCINAE 616
QY 605 CV---SSSYLGVDCTTDKNDRGVGNKKGCHCHCSASYLPDPCSVQSDLWPGGSDSGNFP 662
Db 617 CVDIEKAYKSTNCSS-KCKGHAVCDHELQCCQCEGWIIPDCDDSS----- 660
QY 663 VALPARLPERRYENIYHSPKMRWPFELFIPFIIICVLIAIMVKVNFQKRWRTEDYSS 722
Db 661 -----VVFHFSIVVGVLV---PMAVIF-VVYAMVIRHQSSREKQK-----K 697
QY 723 DEOPES 728
Db 698 DQRPLS 703

Search completed: January 10, 2005, 22:29:19
Job time : 42 secs

```